GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 18:38:27 ; Search time 6228 Seconds
(without alignments)
11467.291 Million cell updates/sec

Title: US-09-973-827-3
Perfect score: 2454
Sequence: 1 cgcggcggaggtgtagtttg.....tctgtatttgtgtatatcat 2454

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 841850

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	27	1.1	27	6	AR090566	AR090566 Sequence
c	2	27	1.1	27	6	AR197601	AR197601 Sequence
	3	26	1.1	26	6	AR090565	AR090565 Sequence
	4	26	1.1	26	6	AR197600	AR197600 Sequence
	5	25	1.0	49	1	ECOINSW	M31752 E.coli inse
c	6	23.2	0.9	50	10	MMU41943	U41943 Mus musculu
c	7	23	0.9	50	6	AX161232	AX161232 Sequence
	8	22.4	0.9	50	6	AX157048	AX157048 Sequence
c	9	22.2	0.9	43	6	AX483407	AX483407 Sequence
c	10	22	0.9	47	6	AX319484	AX319484 Sequence
	11	22	0.9	48	6	AX319485	AX319485 Sequence
	12	21.4	0.9	41	6	AX045523	AX045523 Sequence
	13	21.2	0.9	45	6	A19511	A19511 oligonucleo
	14	21.2	0.9	48	6	AX223683	AX223683 Sequence
c	15	21	0.9	43	6	AX484546	AX484546 Sequence
	16	21	0.9	50	6	AR022597	AR022597 Sequence
	17	21	0.9	50	6	AR037612	AR037612 Sequence
	18	21	0.9	50	6	AR178070	AR178070 Sequence
	19	21	0.9	50	6	AX160550	AX160550 Sequence
c	20	20.8	0.8	43	6	AR166164	AR166164 Sequence
c	21	20.8	0.8	43	6	E33229	E33229 Novel metha
	22	20.8	0.8	45	6	AX320846	AX320846 Sequence
c	23	20.8	0.8	45	6	AX320847	AX320847 Sequence
	24	20.8	0.8	48	6	AX426511	AX426511 Sequence
	25	20.8	0.8	48	9	HSZ95110	Z95110 H.sapiens m
	26	20.8	0.8	49	6	AX456410	AX456410 Sequence
	27	20.6	0.8	36	6	E14547	E14547 DNA probe f
	28	20.6	0.8	40	6	AR149458	AR149458 Sequence
	29	20.6	0.8	40	6	E49430	E49430 Method for
c	30	20.6	0.8	43	6	AX484404	AX484404 Sequence
	31	20.6	0.8	44	6	AR172650	AR172650 Sequence
	32	20.6	0.8	45	9	HUMAIDSB1	M60135 Human/HIV 5
	33	20.6	0.8	48	6	AX223789	AX223789 Sequence
	34	20.6	0.8	48	6	AX426775	AX426775 Sequence
c	35	20.6	0.8	50	6	AR120175	AR120175 Sequence
c	36	20.6	0.8	50	6	AR126173	AR126173 Sequence
c	37	20.6	0.8	50	6	AR177989	AR177989 Sequence
c	38	20.6	0.8	50	6	AX061557	AX061557 Sequence
	39	20.4	0.8	40	6	AR019538	AR019538 Sequence
	40	20.4	0.8	40	6	AR109499	AR109499 Sequence
	41	20.4	0.8	40	6	I55670	I55670 Sequence 23
	42	20.4	0.8	40	6	I76447	I76447 Sequence 23
c	43	20.4	0.8	48	6	AR032406	AR032406 Sequence
c	44	20.4	0.8	48	6	AR209070	AR209070 Sequence
	45	20.4	0.8	48	6	AX223778	AX223778 Sequence

ALIGNMENTS

RESULT 1
AR090566/c
LOCUS AR090566 27 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 686 from patent US 5994076.
ACCESSION AR090566
VERSION AR090566.1 GI:10017321
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 5994076-A 686 30-NOV-1999;
FEATURES Location/Qualifiers

```
     source          1. .27
                     /organism="unknown"
BASE COUNT        4 a      4 c     10 g      9 t
ORIGIN

  Query Match             1.1%;  Score 27;  DB 6;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+05;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     827 ATACCAGATTCGCACAGCACCCACTAG 853
           |||||||||||||||||||||||||||
Db      27 ATACCAGATTCGCACAGCACCCACTAG 1


RESULT 2
AR197601/c
LOCUS       AR197601                  27 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 686 from patent US 6352829.
ACCESSION   AR197601
VERSION     AR197601.1  GI:20247450
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 6352829-A 686 05-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
BASE COUNT        4 a      4 c     10 g      9 t
ORIGIN

  Query Match             1.1%;  Score 27;  DB 6;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+05;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     827 ATACCAGATTCGCACAGCACCCACTAG 853
           |||||||||||||||||||||||||||
Db      27 ATACCAGATTCGCACAGCACCCACTAG 1


RESULT 3
AR090565
LOCUS       AR090565                  26 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 685 from patent US 5994076.
ACCESSION   AR090565
VERSION     AR090565.1  GI:10017320
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 5994076-A 685 30-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        9 a      5 c      9 g      3 t
ORIGIN

  Query Match             1.1%;  Score 26;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+05;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     507 GCACCAGGAGTGCCAAGGATTGAAGA 532
           ||||||||||||||||||||||||||
Db       1 GCACCAGGAGTGCCAAGGATTGAAGA 26


RESULT 4
AR197600
LOCUS       AR197600                  26 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 685 from patent US 6352829.
ACCESSION   AR197600
VERSION     AR197600.1  GI:20247449
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 6352829-A 685 05-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        9 a      5 c      9 g      3 t
ORIGIN

  Query Match             1.1%;  Score 26;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+05;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     507 GCACCAGGAGTGCCAAGGATTGAAGA 532
           ||||||||||||||||||||||||||
Db       1 GCACCAGGAGTGCCAAGGATTGAAGA 26


RESULT 5
ECOINSW
LOCUS       ECOINSW                   49 bp    DNA     linear   BCT 26-APR-1993
DEFINITION  E.coli insertion site for transposon Tn1545.
ACCESSION   M31752
VERSION     M31752.1  GI:146490
KEYWORDS    .
SOURCE      E.coli DNA.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Caillaud,F. and Courvalin,P.
  TITLE     Nucleotide sequence of the ends of the conjugative shuttle
            transposon Tn1545
  JOURNAL   Mol. Gen. Genet. 209 (1), 110-115 (1987)
  MEDLINE   88038347
   PUBMED   2823065
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
BASE COUNT       12 a     10 c      0 g     27 t
ORIGIN

  Query Match             1.0%;  Score 25;  DB 1;  Length 49;
  Best Local Similarity   75.6%;  Pred. No. 1.1e+06;
  Matches   31;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    2300 CAATATTTGGATTGTCATTCTTACAAAACATTTTTTTGTTC 2340
           ||| ||||   || | ||| ||| ||| |||||||||| |||
Db       8 CAACATTTCCTTTTTTATTATTAAAAATCATTTTTTTCTTC 48


RESULT 6
MMU41943/c
LOCUS       MMU41943                  50 bp    DNA     linear   ROD 05-JAN-1996
DEFINITION  Mus musculus recombination between immunoglobulin heavy chain and
            c-myc.
ACCESSION   U41943
VERSION     U41943.1  GI:1147577
KEYWORDS    .
SOURCE      house mouse strain=BALB/cAn.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Muller,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-1995) Jurgen R. Muller, Lab of Genetics, NIH/NCI,
            Bldg. 37, Room 2B09, 37 Convent Dr., Bethesda, MD 20892-4255, USA
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Mus musculus"
                     /strain="BALB/cAn"
                     /db_xref="taxon:10090"
                     /chromosome="T(12;15)"
                     /map="T(12f1;15d2)"
                     /tissue_type="Peyer's patch 7 days post pristane"
                     /dev_stage="7 days post pristane"
BASE COUNT       12 a      11 c      18 g       9 t
ORIGIN

  Query Match             0.9%;  Score 23.2;  DB 10;  Length 50;
  Best Local Similarity   77.8%;  Pred. No. 2.6e+06;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1473 CAGAATTACCCCAGCCTCTTGAGCTGAAGTAATGTG 1508
          ||||  |||||||||||| |||||||||  ||   ||
Db     37 CAGATCTACCCCAGCCTTTTGAGCTGAGCTAGGCTG 2


RESULT 7
AX161232/c
LOCUS       AX161232                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4560 from Patent WO0140521.
ACCESSION   AX161232
VERSION     AX161232.1  GI:14542563
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4560 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43958770"
     misc_feature    26
                     /note="2 of 2 allelic variants (4559 is other entry)"
BASE COUNT       16 a       4 c       5 g      25 t
ORIGIN

  Query Match             0.9%;  Score 23;  DB 6;  Length 50;
  Best Local Similarity   74.4%;  Pred. No. 2.8e+06;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1160 AAAGACAAAATAAACATTTTATTTTCTAAACATTTCTTT 1198
          || |  |||| |||| || |||||| |||| ||||  ||
Db     39 AATGGAAAAAAAAACCTTGTATTTTGTAAATATTTTCTT 1


RESULT 8
AX157048
LOCUS       AX157048                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 376 from Patent WO0140521.
ACCESSION   AX157048
VERSION     AX157048.1  GI:14538379
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 376 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg44921986"
     misc_feature    26
                     /note="2 of 2 allelic variants (375 is other entry)"
BASE COUNT       15 a       3 c       6 g      26 t
ORIGIN

  Query Match             0.9%;  Score 22.4;  DB 6;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 3.7e+06;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2310 ATTGTCATTCTTACAAAACATTTTTTTGTTCTCTTGTAAAAAGAGTAG 2357
          ||| |  |||||| |||| |||||  | || | |  ||   |||| ||
Db      1 ATTTTTTTTCTTAAAAAAATTTTTTGTTTTTTTTAATACTCAGAGGAG 48


RESULT 9
AX483407/c
LOCUS       AX483407                 43 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 707 from Patent WO02053728.
ACCESSION   AX483407
VERSION     AX483407.1  GI:22317827
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 707 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT        6 a      13 c       1 g      23 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 6;  Length 43;
  Best Local Similarity   69.8%;  Pred. No. 4.1e+06;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    152 TGGAGATGCAGCTGTAACAGAAGCTGAAAACCAACAAATGACA 194
          ||||| || | ||| || |||||  ||||   |  |||||| |
Db     43 TGGAGTTGAATCTGGAAAAGAAGAAGAAAGGAAGAAAATGAAA 1


RESULT 10
AX319484/c
LOCUS       AX319484                 47 bp    DNA     linear   PAT 14-DEC-2001
DEFINITION  Sequence 13 from Patent WO0182962.
ACCESSION   AX319484
VERSION     AX319484.1  GI:17901270
KEYWORDS    .
SOURCE      Human immunodeficiency virus.
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
```

```
REFERENCE   1
  AUTHORS   Rovinski,B., Tartaglia,J., Cao,S.X., Persson,R. and Klein,M.H.
  TITLE     Immunizing against hiv infection
  JOURNAL   Patent: WO 0182962-A 13 08-NOV-2001;
            Aventis Pasteur Limited (CA)
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Human immunodeficiency virus"
                     /db_xref="taxon:12721"
BASE COUNT        9 a      15 c       2 g      21 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 47;
  Best Local Similarity   73.7%;  Pred. No. 4.4e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     968 AGAGTGTCGTAGAAAGAAGAAAGAATATGTGAAATGTT 1005
           ||||||  | ||| |||| |||||||||    |||| |
Db      47 AGAGTGGTGCAGAGAGAAAAAAGAATATTCATAATGAT 10


RESULT 11
AX319485
LOCUS       AX319485                 48 bp    DNA     linear   PAT 14-DEC-2001
DEFINITION  Sequence 14 from Patent WO0182962.
ACCESSION   AX319485
VERSION     AX319485.1  GI:17901271
KEYWORDS    .
SOURCE      Human immunodeficiency virus.
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   Rovinski,B., Tartaglia,J., Cao,S.X., Persson,R. and Klein,M.H.
  TITLE     Immunizing against hiv infection
  JOURNAL   Patent: WO 0182962-A 14 08-NOV-2001;
            Aventis Pasteur Limited (CA)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Human immunodeficiency virus"
                     /db_xref="taxon:12721"
BASE COUNT       21 a       3 c      15 g       9 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 48;
  Best Local Similarity   73.7%;  Pred. No. 4.4e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     968 AGAGTGTCGTAGAAAGAAGAAAGAATATGTGAAATGTT 1005
           ||||||  | ||| |||| |||||||||    |||| |
Db       1 AGAGTGGTGCAGAGAGAAAAAAGAATATTCATAATGAT 38


RESULT 12
AX045523
LOCUS       AX045523                 41 bp    DNA     linear   PAT 24-NOV-2000
DEFINITION  Sequence 21 from Patent WO0066742.
ACCESSION   AX045523
VERSION     AX045523.1  GI:11343976
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Rupar,M.J., Donovan,W.P., Chu,C.R., Pease,E., Tan,Y., Slaney,A.C.,
            Malvar,T.M. and Baum,J.A.
  TITLE     Coleopteran-toxic polypeptide compositions and insect-resistant
            transgenic plants
  JOURNAL   Patent: WO 0066742-A 21 09-NOV-2000;
            MONSANTO COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
BASE COUNT       19 a       4 c       5 g      13 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 41;
  Best Local Similarity   71.8%;  Pred. No. 5.9e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    1855 ATTTTAAATTAAAGCAAAGTAAATAAAAGTACAAAGCAT 1893
           || ||| ||  ||  |||||| || ||| | |||| |||
Db       1 ATGTTAGATACAAATAAAGTATATGAAATTTCAAATCAT 39


RESULT 13
A19511
LOCUS       A19511                   45 bp    DNA     linear   PAT 10-JUN-1994
DEFINITION  oligonucleotide 61-83 Hepatititis A virus.
ACCESSION   A19511
VERSION     A19511.1  GI:583227
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Brown,A.L., Clarke,B.E. and Rowlands,D.J.
  TITLE     Chimaeric hepadnavirus core antigen proteins
  JOURNAL   Patent: EP 0421635-A 37 10-APR-1991;
            THE WELLCOME FOUNDATION LIMITED
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       11 a       11 c      13 g      10 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 45;
  Best Local Similarity   69.0%;  Pred. No. 6.3e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy     604 CTAGCAGTGGACAGTATATTGCCATTACCCAGGGAGGAGCAA 645
           |||||| || |||| || || |   | | ||||  ||||| |
Db       2 CTAGCACTGAACAGAATGTTCCGGATCCTCAGGTTGGAGCTA 43


RESULT 14
AX223683
LOCUS       AX223683                 48 bp    mRNA    linear   PAT 07-SEP-2001
DEFINITION  Sequence 9125 from Patent WO0159103.
ACCESSION   AX223683
VERSION     AX223683.1  GI:15551407
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Blatt,L., Mcswiggen,J. and Chowrira,B.M.
  TITLE     Method and reagent for the modulation and diagnosis of cd20 and
            nogo gene expression
  JOURNAL   Patent: WO 0159103-A 9125 16-AUG-2001;
            RIBOZYME PHARMACEUTICALS, INC. (US) ; Blatt, Lawrence (US) ;
            McSwiggen, James (US) ; Chowrira, Bharat M. (US)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Nucleic Acid"
BASE COUNT       12 a       12 c      16 g       8 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 48;
  Best Local Similarity   69.0%;  Pred. No. 6.3e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;
```

```
Qy    443 GCGAAGGGAAATTCTTTCAAGGAGGCCTTCCTACAGGAAAAT 484
          ||| ||| || | | |||||||||   | |||   |||||| |
Db      7 GCGGAGGAAACTCCCTTCAAGGACATCGTCCGGGAGGAAAGT 48


RESULT 15
AX484546/c
LOCUS       AX484546                 43 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 1846 from Patent WO02053728.
ACCESSION   AX484546
VERSION     AX484546.1  GI:22318898
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 1846 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT       12 a      6 c      4 g     21 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 43;
  Best Local Similarity   73.0%;  Pred. No. 7e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1859 TAAATTAAAGCAAAGTAAATAAAAGTACAAAGCATAT 1895
           ||  |||| ||| | |||| ||||| |  |||| |||
Db     41 TAGTTTAACGCATATTAAAAAAAAGAAATAAGCTTAT 5


Search completed: March  9, 2003, 20:52:22
Job time : 6232 secs
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 18:19:02 ; Search time 510 Seconds
(without alignments)
10836.086 Million cell updates/sec

Title: US-09-973-827-3
Perfect score: 2454
Sequence: 1 cgcggcggaggtgtagtttg.....tctgtatttgtgtatatcat 2454

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2166140

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	27	1.1	27	24	ABK66598	Human gene specifi
c	2	26.6	1.1	47	21	AAZ67081	Human map-related
	3	26	1.1	26	24	ABK66597	Human gene specifi
c	4	24	1.0	34	24	ABL40868	CREB cDNA amplifyi
c	5	24	1.0	34	24	ABL40870	CREB cDNA amplifyi
c	6	23.4	1.0	31	19	AAV21495	Plasmid pRSV-CREB3
	7	23.4	1.0	47	21	AAZ66915	Human map-related
c	8	23	0.9	50	22	AAI77619	Human silent SNP c
c	9	22.6	0.9	50	22	AAL34381	Human SNP oligonuc
	10	22.4	0.9	50	22	AAI73435	Human silent SNP c
c	11	22.2	0.9	45	17	AAT07082	Primer ENDO.1, amp
	12	22	0.9	33	24	ABL40867	CREB cDNA amplifyi
c	13	22	0.9	46	24	ABL01753	Human MSH2 (hMSH2)
c	14	22	0.9	47	24	AAI71996	vCP1579 primer HIV
	15	22	0.9	48	24	AAI71997	vCP1579 primer HIV
	16	21.6	0.9	47	21	AAZ67881	Human map-related
c	17	21.6	0.9	49	20	AAX35026	Trailer oligonucle
	18	21.4	0.9	40	13	AAQ25032	Oligonucleotide sp
	19	21.4	0.9	41	21	AAC87847	Bacillus thuringie
	20	21.4	0.9	50	22	AAL28587	Human SNP oligonuc
c	21	21.4	0.9	50	22	AAL28695	Human SNP oligonuc
	22	21.2	0.9	30	20	AAX76677	CREB341 phosphoryl
	23	21.2	0.9	45	12	AAQ11235	Encodes Hepatitis
	24	21.2	0.9	48	23	ABK09125	Human CD20 Amberzy
	25	21.2	0.9	50	22	AAL29818	Human SNP oligonuc
	26	21	0.9	32	19	AAV21501	Plasmid pRSV-CREB3
c	27	21	0.9	33	20	AAX76678	CREB341 phosphoryl
	28	21	0.9	47	21	AAZ69260	Human map-related
	29	21	0.9	50	22	AAI76937	Human silent SNP c
c	30	20.8	0.8	43	21	AAZ61260	Primer BsYqk-G1 fo
c	31	20.8	0.8	45	21	AAZ55674	HaEPV transfer vec
	32	20.8	0.8	47	21	AAZ69043	Human map-related
	33	20.8	0.8	48	24	ABK22200	Human ERG amberzym
	34	20.8	0.8	49	24	ABL99285	Synthetic Renilla
	35	20.8	0.8	50	22	AAL34100	Human SNP oligonuc
c	36	20.8	0.8	50	22	AAL34577	Human SNP oligonuc
c	37	20.6	0.8	36	19	AAV36962	Nucleotide sequenc
	38	20.6	0.8	40	22	AAC83643	Human c-fos strept
	39	20.6	0.8	44	21	AAZ90218	Wild-type phoA pro
c	40	20.6	0.8	44	24	ABL36027	A. thaliana connec
c	41	20.6	0.8	44	24	ABL36030	A. thaliana connec
	42	20.6	0.8	48	23	ABK09231	Human CD20 Amberzy
	43	20.6	0.8	48	24	ABK22464	Human ERG amberzym
	44	20.6	0.8	50	11	AAQ06172	Probe derived from
	45	20.6	0.8	50	14	AAQ37565	RP10 derived BMP-6

ALIGNMENTS

RESULT 1
ABK66598/c
ID ABK66598 standard; DNA; 27 BP.
XX
AC ABK66598;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human gene specific PCR primer #686.
XX
KW Primer; ss; DNA microarray; differential expression analysis; human.
XX
OS Homo sapiens.
XX
PN US6352829-B1.
XX
PD 05-MAR-2002.
XX
PF 05-JAN-1999; 99US-0225928.
XX
PR 21-MAY-1997; 97US-0859998.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Chenchik A, Jokhadze G, Bibilashvilli R;
XX
DR WPI; 2002-314699/35.
XX
PT Producing sub-population of labeled nucleic acids, useful for analysing
PT differences in RNA profiles between several different physiological
PT sources, using set of distinct gene specific primers -
XX

```
PS   Example 3; SEQ ID No 686; 11pp; English.
XX
CC   The invention relates to producing a sub-population of labeled nucleic
CC   acids (NAs) comprising contacting a NA sample from a physiological
CC   source, with a pool of 50 distinct gene specific primers under suitable
CC   conditions to enzymatically generate sub-population of NAs, where
CC   each gene specific primer has a sequence complementary to a distinct
CC   mRNA, and each labeled NA is generated using a single gene specific
CC   primer. The method is useful for producing a sub-population of labeled
CC   NAs which is useful for analysing the differences in the RNA profiles
CC   between several different physiological sources, where the method
CC   comprises producing subpopulation of labeled NAs for the different
CC   physiological sources, comprising the populations for each physiological
CC   source to identify differences in the population, where the comparison
CC   is preferably performed by hybridising the labeled NAs for each of the
CC   distinct physiological sources to an array of probe NAs stably
CC   associated with the surface of a substrate to produce a hybridisation
CC   pattern for each of the sources, and comparing the patterns for each of
CC   the sources, where differential gene expression assays are
CC   utilised in differential expression analysis of diseased a normal
CC   tissue e.g. neoplastic a normal tissue, or different tissue or
CC   subtissue types. The present sequence is a human gene specific PCR
CC   primer used in the method of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from USPTO at
CC   http.wipo.seqdata.uspto.gov/sequence.html?DocID=6352829B1.
XX
SQ   Sequence 27 BP; 4 A; 4 C; 10 G; 9 T; 0 other;

  Query Match             1.1%;  Score 27;  DB 24;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+04;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        827 ATACCAGATTCGCACAGCACCCACTAG 853
              |||||||||||||||||||||||||||
Db         27 ATACCAGATTCGCACAGCACCCACTAG 1


RESULT 2
AAZ67081/c
ID   AAZ67081 standard; DNA; 47 BP.
XX
AC   AAZ67081;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1428.
XX
KW   Human genome; biallelic marker: high density disequilibrium map;
KW   genomic map; haplotype; phenctype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
CS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,T)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082514.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome
XX
PS   Claim 1; Page 535; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 15 A; 4 C; 7 G; 21 T: 0 other;

  Query Match             1.1%;  Score 26.6;  DB 21;  Length 47;
  Best Local Similarity   78.0%;  Pred. No. 2.1e+04;
  Matches   32;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1889 AGCATATTTAGTTAGTACTAAATTCTTAGTAAAATGCTGA 1929
               | ||||||||||| | | |||||| ||| ||||||    |||
Db         43 AACATATTTTAGTAATTGCTAAATACTTTGTAAAACCATGA 3


RESULT 3
ABK66597
ID   ABK66597 standard; DNA; 26 BP.
XX
AC   ABK66597;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Human gene specific PCR primer  #685.
XX
KW   Primer; ss; DNA microarray; differential expression analysis; human.
XX
OS   Homo sapiens.
XX
PN   US6352829-B1.
XX
PD   05-MAR-2002.
XX
PF   05-JAN-1999;   99US-0225928.
XX
PR   21-MAY-1997;   97US-0859998.
XX
PA   (CLON-) CLONTECH LAB INC.
XX
PI   Chenchik A,  Jokhadze G,  Bibilashvilli R;
XX
DR   WPI; 2002-314699/35.
XX
PT   Producing sub-population of labeled nucleic acids, useful for analysing
PT   differences in RNA profiles between several different physiological
PT   sources, using set of distinct gene specific primers
XX
PS   Example 3; SEQ ID No 685; 11pp; English.
XX
CC   The invention relates to producing a sub-population of labeled nucleic
CC   acids (NAs) comprising contacting a NA sample from a physiological
CC   source, with a pool of 50 distinct gene specific primers under suitable
```

CC conditions to enzymatically generate sub-population of NAs, where
CC each gene specific primer has a sequence complementary to a distinct
CC mRNA, and each labeled NA is generated using a single gene specific
CC primer. The method is useful for producing a sub-population of labeled
CC NAs which is useful for analysing the differences in the RNA profiles
CC between several different physiological sources, where the method
CC comprises producing subpopulation of labeled NAs for the different
CC physiological sources, comprising the populations for each physiological
CC source to identify differences in the population, where the comparison
CC is preferably performed by hybridising the labeled NAs for each of the
CC distinct physiological sources to an array of probe NAs stably
CC associated with the surface of a substrate to produce a hybridisation
CC pattern for each of the sources, and comparing the patterns for each of
CC the sources, where differential gene expression assays are
CC utilised in differential expression analysis of diseased a normal
CC tissue e.g. neoplastic a normal tissue, or different tissue or
CC subtissue types. The present sequence is a human gene specific PCR
CC primer used in the method of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from USPTO at
CC http.wipo.seqdata.uspto.gov/sequence.html?DocID=6352829B1.
XX
SQ Sequence 26 BP; 9 A; 5 C; 9 G; 3 T; 0 other;

Query Match 1.1%; Score 26; DB 24; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 507 GCACCAGGAGTGCCAAGGATTGAAGA 532
||||||||||||||||||||||||||
Db 1 GCACCAGGAGTGCCAAGGATTGAAGA 26

RESULT 4
ABL40868/c
ID ABL40868 standard; DNA; 34 BP.
XX
AC ABL40868;
XX
DT 03-JUL-2002 (first entry)
XX
DE CREB cDNA amplifying primer #2.
XX
KW Nucleic acid detection; transcription factor; cytostatic; reporter gene;
KW neuroprotective; cancer diagnosis; Li-Fraumeni syndrome; memory loss;
KW CREB; PCR primer; ss.
XX
OS Synthetic.
XX
PN WO200229043-A1.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-JP08576.
XX
PR 02-OCT-2000; 2000JP-0302100.
XX
PA (HELI-) HELIX RES INST.
XX
PI Oda T, Muramatsu M;
XX
DR WPI; 2002-340106/37.
XX
PT Identifying target gene of transcription factor for isolation in cells
PT with vectors expressing fused protein of transcription factor with its
PT activator and specific reporter gene, useful in cancer diagnosis and
PT therapy -
XX
PS Example 1; Page 12; 32pp; Japanese.
XX
CC The invention relates to a method of detecting the target gene of a
CC specific transcription factor. The method involves (a) supplying cells
CC containing a vector expressing a fused protein of the transcription
CC factor with an activator of this transcription factor and another vector
CC containing a reporter gene; (b) separating cells; and (c) comparing the
CC gene expression dose of the cells. The method is for identifying a target
CC gene of transcription factor for isolation, which is used for application
CC in cancer diagnosis and drug development for e.g. Li-Fraumeni syndrome
CC and long-term memory loss. With expression of the reporter gene as
CC indication, the identification and isolation can be efficiently carried
CC out because its enlarged expression dose in a cell is induced by the
CC transcription factor. Sequences ABL40867-874 represents primers used in
CC the course of the invention.
XX
SQ Sequence 34 BP; 10 A; 4 C; 10 G; 10 T; 0 other;

Query Match 1.0%; Score 24; DB 24; Length 34;
Best Local Similarity 100.0%; Pred. No. 6.8e+04;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1077 CTTTACTGCCACAAATCAGATTAA 1100
||||||||||||||||||||||||
Db 34 CTTTACTGCCACAAATCAGATTAA 11

RESULT 5
ABL40870/c
ID ABL40870 standard; DNA; 34 BP.
XX
AC ABL40870;
XX
DT 03-JUL-2002 (first entry)
XX
DE CREB cDNA amplifying primer #4.
XX
KW Nucleic acid detection; transcription factor; cytostatic; reporter gene;
KW neuroprotective; cancer diagnosis; Li-Fraumeni syndrome; memory loss;
KW CREB; PCR primer; ss.
XX
OS Synthetic.
XX
PN WO200229043-A1.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-JP08576.
XX
PR 02-OCT-2000; 2000JP-0302100.
XX
PA (HELI-) HELIX RES INST.
XX
PI Oda T, Muramatsu M;
XX
DR WPI; 2002-340106/37.
XX
PT Identifying target gene of transcription factor for isolation in cells
PT with vectors expressing fused protein of transcription factor with its
PT activator and specific reporter gene, useful in cancer diagnosis and
PT therapy -
XX
PS Example 1; Page 12; 32pp; Japanese.
XX
CC The invention relates to a method of detecting the target gene of a
CC specific transcription factor. The method involves (a) supplying cells
CC containing a vector expressing a fused protein of the transcription
CC factor with an activator of this transcription factor and another vector
CC containing a reporter gene; (b) separating cells; and (c) comparing the
CC gene expression dose of the cells. The method is for identifying a target
CC gene of transcription factor for isolation, which is used for application
CC in cancer diagnosis and drug development for e.g. Li-Fraumeni syndrome
CC and long-term memory loss. With expression of the reporter gene as
CC indication, the identification and isolation can be efficiently carried
CC out because its enlarged expression dose in a cell is induced by the
CC transcription factor. Sequences ABL40867-874 represents primers used in
CC the course of the invention.

```
XX
SQ   Sequence 34 BP; 10 A; 4 C; 10 G; 10 T; 0 other;

  Query Match             1.0%;  Score 24;  DB 24;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 6.8e+04;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1077 CTTTACTGCCACAAATCAGATTAA 1100
          ||||||||||||||||||||||||
Db     34 CTTTACTGCCACAAATCAGATTAA 11


RESULT 6
AAV21495/c
ID   AAV21495 standard; DNA; 31 BP.
XX
AC   AAV21495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Plasmid pRSV-CREB341 3' CREB 283 primer.
XX
KW   Protein-protein interaction; disruption; binding protein; modulator;
KW   inhibitor; therapeutic; prophylactic; primer; ss.
XX
OS   Synthetic.
XX
PN   WO9813502-A2.
XX
PD   02-APR-1998.
XX
PF   26-SEP-1997;   97WO-US17276.
XX
PR   27-SEP-1996;   96US-0721730.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Goodman RH,  Hoekstra MF;
XX
DR   WPI; 1998-230711/20.
XX
PT   Identifying agents which disrupt protein/protein interactions -
PT   using host cells transformed with DNA to form transcriptional
PT   activating promoter which indirectly affects production of
PT   selectable marker
XX
PS   Example 1; Page 27; 96pp; English.
XX
CC   Oligonucleotides AAV21487 and AAV21503 are used in the construction of
CC   plasmids which are used in a method to identify agents which disrupt
CC   protein/protein interactions using a host cell which is transformed or
CC   transfected with DNA. The host cells are useful to demonstrate in vivo
CC   binding capacity of both known and suspected binding partner proteins
CC   in a recombinant system. The described expression system permits
CC   systematic analysis of the structure and function of a particular
CC   binding protein, thus permitting identification and/or synthesis of
CC   potential modulators of the physiological activity of the binding
CC   proteins. The system can be used to identify and improve molecules which
CC   are capable of inhibiting specific and general protein/protein
CC   interactions. Inhibitors identified by the methods can then be examined
CC   for utility in vivo as therapeutic and/or prophylactic medicaments for
CC   conditions associated with various protein/protein interactions. The
CC   system provides a sensitive assay and provides a positive signal when a
CC   protein/protein interaction is disrupted.
XX
SQ   Sequence 31 BP; 3 A; 10 C; 10 G; 8 T; 0 other;

  Query Match             1.0%;  Score 23.4;  DB 19;  Length 31;
  Best Local Similarity   96.0%;  Pred. No. 8.9e+04;
  Matches   24;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    903 CAGCCTGCTGAAGAAGCAGCACGAA 927
          |||||||||||||||||||||||| |
Db     31 CAGCCTGCTGAAGAAGCAGCACGGA 7


RESULT 7
AAZ66915
ID   AAZ66915 standard; DNA; 47 BP.
XX
AC   AAZ66915;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1262.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,C)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 1; Page 499; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3357, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 23 A; 1 C; 8 G; 15 T; 0 other;

  Query Match             1.0%;  Score 23.4;  DB 21;  Length 47;
  Best Local Similarity   73.2%;  Pred. No. 9.8e+04;
  Matches   30;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1364 AAGAAGTAATAATTTGTTTACTTGTAAATTGATGGGAGAAA 1404
          || | | |||||||||     |||||||||  | |||||||
Db      1 AAAATGGAATAATTTGAACTTTTGTAAATTAGTAGGAGAAA 41
```

```
RESULT 8
AAI77619/c
ID   AAI77619 standard; DNA; 50 BP.
XX
AC   AAI77619;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:4560.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0158138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 1906; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 50 BP; 16 A; 4 C; 5 G; 25 T; 0 other;

  Query Match             0.9%;  Score 23;  DB 22;  Length 50;
  Best Local Similarity   74.4%;  Pred. No. 1.2e+05;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1160 AAAGACAAAATAAACATTTTATTTTCTAAACATTTCTTT 1198
            || |  |||| |||| || |||||| |||| ||||  ||
Db     39 AATGGAAAAAAAAAACCTTGTATTTTGTAAATATTTTCTT 1

RESULT 9
AAL34381/c
ID   AAL34381 standard; DNA; 50 BP.
XX
AC   AAL34381;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #7589.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 3577; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 18 A; 8 C; 11 G; 13 T; 0 other;

  Query Match             0.9%;  Score 22.6;  DB 22;  Length 50;
  Best Local Similarity   68.9%;  Pred. No. 1.5e+05;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2300 CAATATTTGGATTGTCATTCTTACAAAACATTTTTTTGTTCTCTT 2344
          ||||| ||  ||||| || | | ||||| || |  | |  |||||
Db     49 CAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTT 5

RESULT 10
AAI73435
ID   AAI73435 standard; DNA; 50 BP.
XX
AC   AAI73435;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:376.
```

XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 169; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 50 BP; 15 A; 3 C; 6 G; 26 T; 0 other;

```
  Query Match             0.9%;  Score 22.4;  DB 22;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+05;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    2310 ATTGTCATTCTTACAAAACATTTTTTTGTTCTCTTGTAAAAAGAGTAG 2357
           ||| |  |||||| |||| ||||| | || | |  ||   |||| ,|
Db       1 ATTTTTTTTCTTAAAAAAATTTTTTGTTTTTTTTAATACTCAGAGGAG 48
```

RESULT 11
AAT07082/c
ID AAT07082 standard; DNA; 45 BP.
XX
AC AAT07082;
XX
DT 02-JUL-1996 (first entry)
XX
DE Primer ENDO.1, amplifies RNase P conjugate.
XX
KW Primer; RNase P conjugate; amplification; PCR; Endo.P1; Endo.P2;
KW sequence specific cleavage; gene therapy; pathogenic RNA;
KW viral transcript; RNA genome; disease-causing mRNA; oncogene; ss.
XX
OS Synthetic.
XX
PN WO9532283-A1.
XX
PD 30-NOV-1995.
XX
PF 23-MAY-1995; 95WO-US06519.
XX
PR 23-MAY-1994; 94US-0247776.
XX
PA (INDV) UNIV INDIANA FOUND.
XX
PI Frank DH, Harris ME, Pace NR;
XX
DR WPI; 1996-020578/02.
XX
PT RNase P RNA conjugates function as sequence-specific endonuclease(s)
PT - useful for gene therapy and study of, e.g. viral activity in
PT vitro.
XX
PS Example; Page 14; 49pp; English.
XX
CC The sequences given in AAT07078-84 are primers which were used in the
CC construction of the RNase P conjugates of the invention. These
CC primers amplify the Endo.P1 and Endo.P2 genes. The conjugate also
CC includes an oligonucleotide which includes a predetermined sequence
CC complementary to and available for hybridisation with a nucleotide
CC sequence of the target sequence. The conjugates cause sequence
CC specific cleavage of the target oligonucleotide. The conjugates are
CC useful in gene therapy for the selective degradation of pathogenic
CC RNA's such as viral transcripts or RNA genomes, or of disease-
CC causing mRNA's such as products of oncogenes.
XX
SQ Sequence 45 BP; 13 A; 9 C; 9 G; 14 T; 0 other;

```
  Query Match             0.9%;  Score 22.2;  DB 17;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 1.7e+05;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1823 TAACCAACTGAACGACAAAGCATGGTATTTGAATT 1857
           ||||||||||||| | |  |  | ||||| |||||
Db      38 TAACCAACTGAACTATAGTGAGTCGTATTAGAATT 4
```

RESULT 12
ABL40867
ID ABL40867 standard; DNA; 33 BP.
XX
AC ABL40867;
XX
DT 03-JUL-2002 (first entry)
XX
DE CREB cDNA amplifying primer #1.
XX
KW Nucleic acid detection; transcription factor; cytostatic; reporter gene;
KW neuroprotective; cancer diagnosis; Li-Fraumeni syndrome; memory loss;
KW CREB; PCR primer; ss.
XX
OS Synthetic.
XX
PN WO200229043-A1.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-JP08576.
XX
PR 02-OCT-2000; 2000JP-0302100.
XX
PA (HELI-) HELIX RES INST.
XX
PI Oda T, Muramatsu M;
XX
DR WPI; 2002-340106/37.
XX

PT Identifying target gene of transcription factor for isolation in cells
PT with vectors expressing fused protein of transcription factor with its
PT activator and specific reporter gene, useful in cancer diagnosis and
PT therapy -
XX
PS Example 1; Page 12; 32pp; Japanese.
XX
CC The invention relates to a method of detecting the target gene of a
CC specific transcription factor. The method involves (a) supplying cells
CC containing a vector expressing a fused protein of the transcription
CC factor with an activator of this transcription factor and another vector
CC containing a reporter gene; (b) separating cells; and (c) comparing the
CC gene expression dose of the cells. The method is for identifying a target
CC gene of transcription factor for isolation, which is used for application
CC in cancer diagnosis and drug development for e.g. Li-Fraumeni syndrome
CC and long-term memory loss. With expression of the reporter gene as
CC indication, the identification and isolation can be efficiently carried
CC out because its enlarged expression dose in a cell is induced by the
CC transcription factor. Sequences ABL40867-874 represents primers used in
CC the course of the invention.
XX
SQ Sequence 33 BP; 10 A; 6 C; 8 G; 9 T; 0 other;

Query Match 0.9%; Score 22; DB 24; Length 33;
Best Local Similarity 83.3%; Pred. No. 1.8e+05;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 110 TAACTAAATGACCATGGAATCTGGAGCCGA 139
| | | ||||||||||||||||||||||| ||
Db 3 TGAATTCATGACCATGGAATCTGGAGCAGA 32

RESULT 13
ABL01753/c
ID ABL01753 standard; DNA; 46 BP.
XX
AC ABL01753;
XX
DT 18-MAR-2002 (first entry)
XX
DE Human MSH2 (hMSH2) intronic sequence SEQ ID NO:106.
XX
KW Human; MLH1; MSH2; hMLH1; hMSH2; variant gene; diagnosis; HNPCC;
KW hereditary non-polyposis colorectal cancer; ds.
XX
OS Homo sapiens.
XX
PN US2001044936-A1.
XX
PD 22-NOV-2001.
XX
PF 22-OCT-1999; 99US-0426548.
XX
PR 22-OCT-1998; 98US-105160P.
XX
PA (ROBB/) ROBBINS D.
PA (LING/) LIN-GOERKE J L.
PA (LING/) LING J C.
XX
PI Robbins D, Lin-Goerke JL, Ling JC;
XX
DR WPI; 2002-105577/14.
XX
PT New variants of the human MLH1 and MSH2 genes for diagnosing or
PT determining a predisposition for hereditary non-polyposis colorectal
PT cancer -
XX
PS Disclosure; Page 4; 38pp; English.
XX
CC The present invention describes a variant human MLH1 or MSH2 gene.
CC Also described are: (1) a method for diagnosing or predicting
CC susceptibility to hereditary non-polyposis colorectal cancer (HNPCC),
CC comprising screening a DNA sample for the variant MLH1 or MSH2 gene
CC where presence of the variant indicates presence of, or susceptibility
CC to HNPCC; (2) a method of identifying mutants in splice donor or
CC acceptor sites of a human MLH1 gene, comprising sequencing splice donor
CC or acceptor sites of the gene with intronic primers for the human MLH1
CC gene and analysing the sequence to identify any mutants; (3) a method of
CC identifying mutants in splice donor or acceptor sites of a human MSH2
CC gene, comprising sequencing splice donor or acceptor sites of the gene
CC with intronic primers for the human MSH2 gene and analysing the sequence
CC to identify any mutants; and (4) a transgenic model system for
CC colorectal cancer comprising cells expressing the variant MLH1 or MSH2
CC gene. The hMLH1 and hMSH2 variants are used to diagnose or determine a
CC patients susceptibility to hereditary non-polyposis colorectal cancer.
CC ABL01648 to ABL01745 and ABL01746 to ABL01831 represent hMLH1 and hMSH2
CC gene fragments from the present invention. ABL01832 to ABL01839
CC represent mutagenic primers used in the exemplification of the present
CC invention.
XX
SQ Sequence 46 BP; 14 A; 5 C; 4 G; 23 T; 0 other;

Query Match 0.9%; Score 22; DB 24; Length 46;
Best Local Similarity 73.7%; Pred. No. 1.9e+05;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1870 AAAGTAAATAAAAGTACAAAGCATATTTTAGTTAGTAC 1907
||| ||| |||||||||| || |||| || || |||
Db 43 AAAAAAAAAAAAAGTACATTACAGATTTGAGATATTAC 6

RESULT 14
AAI71996/c
ID AAI71996 standard; DNA; 47 BP.
XX
AC AAI71996;
XX
DT 21-FEB-2002 (first entry)
XX
DE vCP1579 primer HIVP100.
XX
KW Immunisation; Human Immunodeficiency Virus; HIV; envelope glycoprotein;
KW priming antigen; booster antigen; attenuated; viral vector; primer; PCR;
KW long terminal repeat; LTR; HIV-1; polymerase chain reaction; amplify; ss.
XX
OS Synthetic.
XX
PN WO200182962-A2.
XX
PD 08-NOV-2001.
XX
PF 25-APR-2001; 2001WO-CA00577.
XX
PR 27-APR-2000; 2000US-200011P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Rovinski B, Tartaglia J, Cao S, Persson R, Klein MH;
XX
DR WPI; 2002-034490/04.
XX
PT Immunizing against Human Immunodeficiency Virus (HIV) using primary and
PT booster antigens -
XX
PS Example 4; Page 16; 38pp; English.
XX
CC The sequences given in AAI71986-97 are primers which were used in
CC the production of the recombinant poxvirus, vCP1579. vCP1579 contains
CC the HIV-1 gag and protease genes derived from the HIV-1 isolate IIIB,
CC the gp120 envelope sequences derived from the HIV-1 BxO8 isolate, and
CC sequences encoding a polypeptide encompassing the known human
CC cytotoxic T lymphocytes (CTL) epitopes from HIV-1 NEf and Pol. vCP1579
CC may be used to immunise against Human Immunodeficiency Virus (HIV). The
CC method of the invention for immunising against HIV infection uses priming
CC (DNA encoding an envelope glycoprotein of a primary HIV-1 isolate) and
CC booster antigens (non-infectious, non-replicating immunogenic HIV-like

CC particle comprising the env glycoprotein of primary HIV-1 isolate or
CC an attenuated viral vector that expresses the glycoprotein). The
CC attenuated viral vector vCP1579 encodes the boosting antigen and is
CC deficient in long terminal repeats (LTR's). The methods and vectors
CC are used to immunize patients against HIV-1 infection.
XX
SQ Sequence 47 BP; 9 A; 15 C; 2 G; 21 T; 0 other;

Query Match 0.9%; Score 22; DB 24; Length 47;
Best Local Similarity 73.7%; Pred. No. 1.9e+05;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy     968 AGAGTGTCGTAGAAAGAAGAAAGAATATGTGAAATGTT 1005
           ||||||  | ||| |||| |||||||||    |||| |
Db      47 AGAGTGGTGCAGAGAGAAAAAAGAATATTCATAATGAT 10
```

RESULT 15
AAI71997
ID AAI71997 standard; DNA; 48 BP.
XX
AC AAI71997;
XX
DT 21-FEB-2002 (first entry)
XX
DE vCP1579 primer HIVP99.
XX
KW Immunisation; Human Immunodeficiency Virus; HIV; envelope glycoprotein;
KW priming antigen; booster antigen; attenuated; viral vector; primer; PCR;
KW long terminal repeat; LTR; HIV-1; polymerase chain reaction; amplify; ss.
XX
OS Synthetic.
XX
PN WO200182962-A2.
XX
PD 08-NOV-2001.
XX
PF 25-APR-2001; 2001WO-CA00577.
XX
PR 27-APR-2000; 2000US-200011P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Rovinski B, Tartaglia J, Cao S, Persson R, Klein MH;
XX
DR WPI; 2002-034490/04.
XX
PT Immunizing against Human Immunodeficiency Virus (HIV) using primary and
PT booster antigens -
XX
PS Example 4; Page 15; 38pp; English.
XX
CC The sequences given in AAI71986-97 are primers which were used in
CC the production of the recombinant poxvirus, vCP1579. vCP1579 contains
CC the HIV-1 gag and protease genes derived from the HIV-1 isolate IIIB,
CC the gp120 envelope sequences derived from the HIV-1 Bx08 isolate, and
CC sequences encoding a polypeptide encompassing the known human
CC cytotoxic T lymphocytes (CTL) epitopes from HIV-1 NEf and Pol. vCP1579
CC may be used to immunise against Human Immunodeficiency Virus (HIV). The
CC method of the invention for immunising against HIV infection uses priming
CC (DNA encoding an envelope glycoprotein of a primary HIV-1 isolate) and
CC booster antigens (non-infectious, non-replicating immunogenic HIV-like
CC particle comprising the env glycoprotein of primary HIV-1 isolate or
CC an attenuated viral vector that expresses the glycoprotein). The
CC attenuated viral vector vCP1579 encodes the boosting antigen and is
CC deficient in long terminal repeats (LTR's). The methods and vectors
CC are used to immunize patients against HIV-1 infection.
XX
SQ Sequence 48 BP; 21 A; 3 C; 15 G; 9 T; 0 other;

Query Match 0.9%; Score 22; DB 24; Length 48;
Best Local Similarity 73.7%; Pred. No. 1.9e+05;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy     968 AGAGTGTCGTAGAAAGAAGAAAGAATATGTGAAATGTT 1005
           ||||||  | ||| |||| |||||||||    |||| |
Db       1 AGAGTGGTGCAGAGAGAAAAAAGAATATTCATAATGAT 38
```

Search completed: March 9, 2003, 19:08:20
Job time : 513 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 19:00:11 ; Search time 104 Seconds
(without alignments)
7236.392 Million cell updates/sec

Title: US-09-973-827-3
Perfect score: 2454
Sequence: 1 cgcggcggaggtgtagtttg.....tctgtatttgtgtatatcat 2454

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 609818

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	27	1.1	27	2	US-08-859-998-686	Sequence 686, App
c	2	27	1.1	27	4	US-09-225-928-686	Sequence 686, App
	3	26	1.1	26	2	US-08-859-998-685	Sequence 685, App
	4	26	1.1	26	4	US-09-225-928-685	Sequence 685, App
	5	21	0.9	40	3	US-08-881-094-33	Sequence 33, Appl
	6	21	0.9	50	1	US-07-977-696C-60	Sequence 60, Appl
	7	21	0.9	50	1	US-08-129-930B-60	Sequence 60, Appl
	8	21	0.9	50	4	US-08-976-288A-60	Sequence 60, Appl
c	9	20.8	0.8	43	4	US-09-387-800-5	Sequence 5, Appli
	10	20.6	0.8	40	4	US-09-476-256-29	Sequence 29, Appl
	11	20.6	0.8	44	4	US-09-349-644-13	Sequence 13, Appl
c	12	20.6	0.8	50	2	US-08-980-071-21	Sequence 21, Appl
c	13	20.6	0.8	50	2	US-08-757-536-21	Sequence 21, Appl
c	14	20.6	0.8	50	3	US-09-314-093-21	Sequence 21, Appl
c	15	20.6	0.8	50	3	US-09-250-848-21	Sequence 21, Appl
c	16	20.6	0.8	50	4	US-09-251-885-21	Sequence 21, Appl
c	17	20.6	0.8	50	4	US-09-337-635-21	Sequence 21, Appl
c	18	20.6	0.8	50	4	US-09-337-280-21	Sequence 21, Appl
	19	20.6	0.8	50	6	5187076-16	Patent No. 5187076
	20	20.4	0.8	40	1	US-07-741-940-23	Sequence 23, Appl
	21	20.4	0.8	40	1	US-08-289-548A-23	Sequence 23, Appl
	22	20.4	0.8	40	1	US-08-452-654-23	Sequence 23, Appl
	23	20.4	0.8	40	1	US-08-452-655B-23	Sequence 23, Appl
	24	20.4	0.8	40	3	US-08-450-582-23	Sequence 23, Appl
	25	20.4	0.8	40	4	US-08-449-731-23	Sequence 23, Appl
	26	20.4	0.8	47	4	US-09-338-907-206	Sequence 206, App
	27	20.4	0.8	47	4	US-09-218-207-206	Sequence 206, App
c	28	20.4	0.8	48	1	US-08-171-389-18	Sequence 18, Appl
c	29	20.4	0.8	48	1	US-08-123-936-18	Sequence 18, Appl
c	30	20.4	0.8	48	2	US-08-475-228A-18	Sequence 18, Appl
c	31	20.4	0.8	48	3	US-08-482-080A-18	Sequence 18, Appl
c	32	20.4	0.8	48	4	US-09-354-947-18	Sequence 18, Appl
c	33	20.4	0.8	48	5	PCT-US93-12388-18	Sequence 18, Appl
	34	20	0.8	30	2	US-08-343-443B-113	Sequence 113, App
	35	20	0.8	40	1	US-08-411-795B-170	Sequence 170, App
	36	20	0.8	40	1	US-08-411-795B-172	Sequence 172, App
	37	20	0.8	40	1	US-08-469-319A-170	Sequence 170, App
	38	20	0.8	40	1	US-08-469-319A-172	Sequence 172, App
	39	20	0.8	40	4	US-08-764-114-170	Sequence 170, App
	40	20	0.8	40	4	US-08-764-114-172	Sequence 172, App
	41	20	0.8	40	4	US-08-469-419-170	Sequence 170, App
	42	20	0.8	40	4	US-08-469-419-172	Sequence 172, App
c	43	20	0.8	45	1	US-08-411-795B-171	Sequence 171, App
c	44	20	0.8	45	1	US-08-411-795B-173	Sequence 173, App
c	45	20	0.8	45	1	US-08-411-795B-235	Sequence 235, App

ALIGNMENTS

RESULT 1
US-08-859-998-686/c
; Sequence 686, Application US/08859998
; Patent No. 5994076
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jokhadze, George
; APPLICANT: Bibilashvilli, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,998
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 686:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: oligonucleotide primer
US-08-859-998-686

Query Match 1.1%; Score 27; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 4e+02;

```
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     827 ATACCAGATTCGCACAGCACCCACTAG 853
           ||||||||||||||||||||||||||||
Db      27 ATACCAGATTCGCACAGCACCCACTAG 1


RESULT 2
US-09-225-928-686/c
; Sequence 686, Application US/09225928
; Patent No. 6352829
;   GENERAL INFORMATION:
;        APPLICANT: Chenchik, Alex
;                   Jokhadze, George
;                   Bibilashvilli, Robert
;        TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
;                            EXPRESSION
;        NUMBER OF SEQUENCES: 1375
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson, P.C.
;             STREET: 2200 Sand Hill Road, Suite 100
;             CITY: Menlo Park
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/225,928
;             FILING DATE: 05-Jan-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/859,998
;             FILING DATE: 21-MAY-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Field, Bret E.
;             REGISTRATION NUMBER: 37,620
;             REFERENCE/DOCKET NUMBER: 09096/002001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-322-5070
;             TELEFAX: 415-854-0875
;   INFORMATION FOR SEQ ID NO: 686:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 27 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        FEATURE:
;             OTHER INFORMATION:  oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 686:
US-09-225-928-686

  Query Match             1.1%;  Score 27;  DB 4;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 4e+02;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     827 ATACCAGATTCGCACAGCACCCACTAG 853
           ||||||||||||||||||||||||||||
Db      27 ATACCAGATTCGCACAGCACCCACTAG 1


RESULT 3
US-08-859-998-685
; Sequence 685, Application US/08859998
; Patent No. 5994076
;  GENERAL INFORMATION:
;    APPLICANT:  Chenchik, Alex
;    APPLICANT:  Jokhadze, George
;    APPLICANT:  Bibilashvilli, Robert
;    TITLE OF INVENTION:  METHOD OF ASSAYING DIFFERENTIAL
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1375
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  2200 Sand Hill Road, Suite 100
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/859,998
;      FILING DATE:  21-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Field, Bret E.
;      REGISTRATION NUMBER:  37,620
;      REFERENCE/DOCKET NUMBER:  09096/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  685:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      OTHER INFORMATION:  oligonucleotide primer
US-08-859-998-685

  Query Match             1.1%;  Score 26;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 7e+02;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     507 GCACCAGGAGTGCCAAGGATTGAAGA 532
           ||||||||||||||||||||||||||
Db       1 GCACCAGGAGTGCCAAGGATTGAAGA 26


RESULT 4
US-09-225-928-685
; Sequence 685, Application US/09225928
; Patent No. 6352829
;   GENERAL INFORMATION:
;        APPLICANT: Chenchik, Alex
;                   Jokhadze, George
;                   Bibilashvilli, Robert
;        TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
;                            EXPRESSION
;        NUMBER OF SEQUENCES: 1375
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson, P.C.
;             STREET: 2200 Sand Hill Road, Suite 100
;             CITY: Menlo Park
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
```

```
;             APPLICATION NUMBER: US/09/225,928
;             FILING DATE: 05-Jan-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/859,998
;             FILING DATE: 21-MAY-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Field, Bret E.
;             REGISTRATION NUMBER: 37,620
;             REFERENCE/DOCKET NUMBER: 09096/002001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-322-5070
;             TELEFAX: 415-854-0875
;   INFORMATION FOR SEQ ID NO: 685:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 26 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        FEATURE:
;             OTHER INFORMATION: oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 685:
US-09-225-928-685

  Query Match             1.1%;  Score 26;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 7e+02;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     507 GCACCAGGAGTGCCAAGGATTGAAGA 532
           ||||||||||||||||||||||||||
Db       1 GCACCAGGAGTGCCAAGGATTGAAGA 26


RESULT 5
US-08-881-094-33
; Sequence 33, Application US/08881094A
; Patent No. 6022739
; GENERAL INFORMATION:
;  APPLICANT: Ryan, Clarence A
;  APPLICANT: Pearce, Gregory L
;  APPLICANT: McGurl, Barry F
;  TITLE OF INVENTION: Systemin
;  FILE REFERENCE: 7555-000001CPB
;  CURRENT APPLICATION NUMBER: US/08/881,094A
;  CURRENT FILING DATE: 1997-07-09
;  EARLIER APPLICATION NUMBER: 08/308,887
;  EARLIER FILING DATE: 1994-09-19
;  EARLIER APPLICATION NUMBER: PCT/US93/02428
;  EARLIER FILING DATE: 1993-03-18
;  EARLIER APPLICATION NUMBER: 07/885,412
;  EARLIER FILING DATE: 1992-03-19
;  EARLIER APPLICATION NUMBER: 07/528,956
;  EARLIER FILING DATE: 1990-05-25
;  EARLIER APPLICATION NUMBER: PCT/US91/03685
;  EARLIER FILING DATE: 1991-05-24
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Lycopersicon esculentum
US-08-881-094-33

  Query Match             0.9%;  Score 21;  DB 3;  Length 40;
  Best Local Similarity   73.0%;  Pred. No. 1.5e+04;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     503 TGATGCACCAGGAGTGCCAAGGATTGAAGAAGAGAAG 539
           |||| ||| || | | |||| ||| ||| | ||| ||
Db       4 TGATACACAAGAAATACCAAAGATGGAACATGAGGAG 40


RESULT 6
US-07-977-696C-60
; Sequence 60, Application US/07977696C
; Patent No. 5792852
; GENERAL INFORMATION:
;    APPLICANT:  do Couto, Fernando J.R.
;    APPLICANT:  Ceriani Dr., Roberto L.
;    APPLICANT:  Peterson Dr., Jerry A.
;    APPLICANT:  Padlan Dr., Eduardo A.
;    TITLE OF INVENTION:  Analogue Peptides with Specificity
;    TITLE OF INVENTION:  for Carcinomas and Kit and Diagnostic Vaccination
;    TITLE OF INVENTION:  and Therapeutic Methods.
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 5.0
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/977,696C
;      FILING DATE:  11-16-92
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amzel Ph.D., Viviana
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38227
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 748-6868
;      TELEFAX:  (510) 748-6688
;      TELEX:  n.a.
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-977-696C-60

  Query Match             0.9%;  Score 21;  DB 1;  Length 50;
  Best Local Similarity   73.0%;  Pred. No. 1.6e+04;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     953 CAGGGAAGCAGCTCGAGAGTGTCGTAGAAAGAAGAAA 989
           ||||||||  ||| ||| | || |  ||||  |||||
Db      14 CAGGGAAGCGGCTTGAGTGGGTTGCTGAAATTAGAAA 50


RESULT 7
US-08-129-93CB-60
; Sequence 60, Application US/08129930B
; Patent No. 5804167
; GENERAL INFORMATION:
;    APPLICANT:  do Couto Dr., Fernando J.R.
;    APPLICANT:  Ceriani Dr., Roberto L.
;    APPLICANT:  Peterson Dr., Jerry A.
;    APPLICANT:  Padlan Dr., Eduardo A.
;    TITLE OF INVENTION:  Analogue Peptides With Broad
;    TITLE OF INVENTION:  Carcinoma Specificity, and Kit and
;    TITLE OF INVENTION:  Diagnostic Vaccination and
;    TITLE OF INVENTION:  Therapeutic Methods
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  V. AMZEL & ASSOC.
;      STREET:  2055 No. 5804187th Broadway, Suite 201
;      CITY:  Walnut Creek
```

```
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94596
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 5.0
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/129,930B
;      FILING DATE:  September 30, 1993
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amzel Ph.D., Viviana
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  CRFCC-008A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 521-1333
;      TELEFAX:  (510) 521-3541
;      TELEX:  n.a.
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-129-930B-60

  Query Match             0.9%;  Score 21;  DB 1;  Length 50;
  Best Local Similarity   73.0%;  Pred. No. 1.6e+04;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     953 CAGGGAAGCAGCTCGAGAGTGTCGTAGAAAGAAGAAA 989
           ||||||||  ||| ||| | || |  ||||  |||||
Db      14 CAGGGAAGGGGCTTGAGTGGGTTGCTGAAATTAGAAA 50


RESULT 8
US-08-976-288A-60
; Sequence 60, Application US/08976288A
; Patent No. 6315997
;  GENERAL INFORMATION:
;    APPLICANT:  do Couto Dr., Fernando J.R.
;    APPLICANT:  Ceriani Dr., Roberto L.
;    APPLICANT:  Peterson Dr., Jerry A.
;    APPLICANT:  Padlan Dr., Eduardo A.
;    TITLE OF INVENTION:  Analogue Peptides With Broad
;    TITLE OF INVENTION:  Carcinoma Specificity, and Kit and
;    TITLE OF INVENTION:  Diagnostic Vaccination and
;    TITLE OF INVENTION:  Therapeutic Methods
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder & Poplawski
;      STREET:  444 South Flower St., 19th Floor
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 5.0
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/976,288A
;      FILING DATE:  No. 6315997ember 21, 1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/129,930
;      FILING DATE:  September 30, 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/977,696
;      FILING DATE:  No. 6315997ember 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel Ph.D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P6639938
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;      TELEX:  n.a.
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-976-288A-60

  Query Match             0.9%;  Score 21;  DB 4;  Length 50;
  Best Local Similarity   73.0%;  Pred. No. 1.6e+04;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     953 CAGGGAAGCAGCTCGAGAGTGTCGTAGAAAGAAGAAA 989
           ||||||||  ||| ||| | || |  ||||  |||||
Db      14 CAGGGAAGGGGCTTGAGTGGGTTGCTGAAATTAGAAA 50


RESULT 9
US-09-387-800-5/c
; Sequence 5, Application US/09387800
; Patent No. 6280972
; GENERAL INFORMATION:
;  APPLICANT: YASUEDA, Hisashi
;  TITLE OF INVENTION: NOVEL ACTIVATOR FOR ALCOHOL DEHYDROGENASE AND GENE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: 0010-1036-0
;  CURRENT APPLICATION NUMBER: US/09/387,800
;  CURRENT FILING DATE: 1999-09-01
;  EARLIER APPLICATION NUMBER: JP 10-248297
;  EARLIER FILING DATE: 1998-09-02
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;  LENGTH: 43
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:BS-YQKG1-1
US-09-387-800-5

  Query Match             0.8%;  Score 20.8;  DB 4;  Length 43;
  Best Local Similarity   70.0%;  Pred. No. 1.7e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    1171 AAACATTTTATTTTCTAAACATTTCTTTTTTTCTATGCGC 1210
           ||   |||| | ||||||  ||||| |||| || || | |
Db      40 AATTGTTTTTTCTTCTAATGATTTCATTTTATCGATTCCC 1


RESULT 10
US-09-476-256-29
; Sequence 29, Application US/09476256
; Patent No. 6228592
; GENERAL INFORMATION:
;  APPLICANT: Laboratory of Molecular Biophotonics
;  TITLE OF INVENTION: Nucleic Acid Detection in Cytoplasm
;  FILE REFERENCE: BBP99-02
;  CURRENT APPLICATION NUMBER: US/09/476,256
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 29
;  LENGTH: 40
;  TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: probe
US-09-476-256-29

  Query Match             0.8%;  Score 20.6;  DB 4;  Length 40;
  Best Local Similarity   74.3%;  Pred. No. 1.9e+04;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   1329 GAAGAGACTTCTGCTTTTCAACCCCCACCCTCCTC 1363
          || ||||  |||||||| |  ||||| ||| || |
Db      2 GATGAGAAGTCTGCTTTGCCCCCCCCCCCCCCCCC 36


RESULT 11
US-09-349-644-13
; Sequence 13, Application US/09349644A
; Patent No. 6303340
; GENERAL INFORMATION:
;  APPLICANT: Pollitt, Stephen N.
;  APPLICANT: Buckley, Douglas I.
;  APPLICANT: Stathis, Peter A.
;  APPLICANT: Hartman, Taymar E.
;  TITLE OF INVENTION: METHOD FOR PRODUCING A PEPTIDE WITH A PI
;  TITLE OF INVENTION: ABOVE 8 OR BELOW 5
;  FILE REFERENCE: SCIOS.019A
;  CURRENT APPLICATION NUMBER: US/09/349,644A
;  CURRENT FILING DATE: 1999-07-08
;  EARLIER APPLICATION NUMBER: 60/092,423
;  EARLIER FILING DATE: 1998-07-10
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: E. coli
US-09-349-644-13

  Query Match             0.8%;  Score 20.6;  DB 4;  Length 44;
  Best Local Similarity   74.3%;  Pred. No. 2e+04;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   2112 TATATATTTGTTTATGAAATGTTACATGTAGAAAA 2146
          | | ||||| || ||| | |  |||||| ||||||
Db      3 TTTTTATTTTTTAATGTATTTGTACATGGAGAAAA 37


RESULT 12
US-08-980-071-21/c
; Sequence 21, Application US/08980071
; Patent No. 5914318
;  GENERAL INFORMATION:
;    APPLICANT:  Baum, James A.
;    APPLICANT:  Gilmer, Amy Jelen
;    APPLICANT:  Mettus, Anne-Marie Light
;    TITLE OF INVENTION:  TRANSGENIC PLANTS EXPRESSING
;    TITLE OF INVENTION:  LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/980,071
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/757,536
;      FILING DATE:  27-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  MECO:206
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-980-071-21

  Query Match             0.8%;  Score 20.6;  DB 2;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 2.1e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1332 GAGACTTCTGCTTTTCAACCCCCACCCTCCTCAAGAAGTAATA 1374
          || | || |  || |||||||||| || ||| | || | | |
Db     50 GACATTTATCGTTGTCAACCCCCATGCTTCTCCAAAAATTACA 8


RESULT 13
US-08-757-536-21/c
; Sequence 21, Application US/08757536
; Patent No. 5942664
;  GENERAL INFORMATION:
;    APPLICANT:  Baum, James A.
;    APPLICANT:  Gilmer, Amy Jelen
;    APPLICANT:  Mettus, Anne-Marie Light
;    TITLE OF INVENTION:  Bacillus thuringiensis CryIC
;    TITLE OF INVENTION:  Compositions Toxic to Lepidopteran Insects and Methods fc
;    TITLE OF INVENTION:  Making CryIC Mutants
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White and Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,536
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  MOBT:023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-757-536-21

  Query Match             0.8%;  Score 20.6;  DB 2;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 2.1e+04;
```

```
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1332 GAGACTTCTGCTTTTCAACCCCCACCCTCCTCAAGAAGTAATA 1374
          || | || |  || ||||||||||  || ||| | || | | |
Db     50 GACATTTATCGTTGTCAACCCCCATGCTTCTCCAAAAATTACA 8


RESULT 14
US-09-314-093-21/c
; Sequence 21, Application US/09314093
; Patent No. 6033874
;  GENERAL INFORMATION:
;    APPLICANT:  Baum, James A.
;    APPLICANT:  Gilmer, Amy Jelen
;    APPLICANT:  Mettus, Anne-Marie Light
;    TITLE OF INVENTION:  TRANSGENIC PLANTS EXPRESSING
;    TITLE OF INVENTION:  LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/314,093
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/980,071
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  MECO:205
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-314-093-21

  Query Match             0.8%;  Score 20.6;  DB 3;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 2.1e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1332 GAGACTTCTGCTTTTCAACCCCCACCCTCCTCAAGAAGTAATA 1374
          || | || |  || ||||||||||  || ||| | || | | |
Db     50 GACATTTATCGTTGTCAACCCCCATGCTTCTCCAAAAATTACA 8


RESULT 15
US-09-250-848-21/c
; Sequence 21, Application US/09250848
; Patent No. 6153814
;  GENERAL INFORMATION:
;    APPLICANT:  Baum, James A.
;    APPLICANT:  Gilmer, Amy Jelen
;    APPLICANT:  Mettus, Anne-Marie Light
;    TITLE OF INVENTION:  Bacillus thuringiensis Cry1C
;    TITLE OF INVENTION:  Compositions Toxic to Lepidopteran Insects and Methods for
;    TITLE OF INVENTION:  Making Cry1C Mutants
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White and Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/250,848
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,536
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  MOBT:023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-250-848-21

  Query Match             0.8%;  Score 20.6;  DB 3;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 2.1e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1332 GAGACTTCTGCTTTTCAACCCCCACCCTCCTCAAGAAGTAATA 1374
          || | || |  || ||||||||||  || ||| | || | | |
Db     50 GACATTTATCGTTGTCAACCCCCATGCTTCTCCAAAAATTACA 8


Search completed: March  9, 2003, 21:46:21
Job time : 106 secs
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 19:08:27 ; Search time 197 Seconds
(without alignments)
8320.180 Million cell updates/sec

Title: US-09-973-827-3
Perfect score: 2454
Sequence: 1 cgcggcggaggtgtagtttg.....tctgtatttgtgtatatcat 2454

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 261222

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	22	0.9	46	10	US-09-426-548-106	Sequence 106, App
c	2	22	0.9	47	10	US-09-842-883-13	Sequence 13, Appl
	3	22	0.9	48	10	US-09-842-883-14	Sequence 14, Appl
	4	20.8	0.8	45	10	US-09-838-386-17	Sequence 17, Appl
c	5	20.8	0.8	45	10	US-09-838-386-18	Sequence 18, Appl
c	6	20.4	0.8	43	9	US-10-027-806-92	Sequence 92, Appl
c	7	20.4	0.8	43	9	US-10-034-623-92	Sequence 92, Appl
	8	20.4	0.8	47	9	US-09-853-526-206	Sequence 206, App
	9	20.4	0.8	47	10	US-09-901-484A-206	Sequence 206, App
c	10	20.2	0.8	44	9	US-10-079-623-71	Sequence 71, Appl
	11	20	0.8	46	10	US-09-827-289-22	Sequence 22, Appl
	12	20	0.8	48	9	US-09-864-785-3523	Sequence 3523, Ap
c	13	19.8	0.8	41	10	US-09-729-402-28	Sequence 28, Appl
	14	19.8	0.8	45	10	US-09-921-398-18	Sequence 18, Appl
	15	19.8	0.8	45	10	US-09-263-959-486	Sequence 486, App
c	16	19.8	0.8	46	9	US-09-792-793A-48	Sequence 48, Appl
	17	19.8	0.8	48	9	US-09-825-805-55	Sequence 55, Appl
	18	19.6	0.8	36	10	US-09-062-113-50	Sequence 50, Appl
	19	19.6	0.8	45	10	US-09-910-059-127	Sequence 127, App
	20	19.6	0.8	48	9	US-09-864-785-3014	Sequence 3014, Ap
	21	19.4	0.8	38	9	US-09-864-785-1089	Sequence 1089, Ap
	22	19.4	0.8	48	9	US-09-864-785-3448	Sequence 3448, Ap
	23	19.4	0.8	48	9	US-09-864-785-3573	Sequence 3573, Ap
c	24	19	0.8	31	10	US-09-801-274-399	Sequence 399, App
c	25	19	0.8	35	9	US-09-904-557-9	Sequence 9, Appli
	26	19	0.8	46	10	US-09-827-289-26	Sequence 26, Appl
	27	19	0.8	48	9	US-09-864-785-3078	Sequence 3078, Ap
	28	19	0.8	48	9	US-09-864-785-3557	Sequence 3557, Ap
	29	19	0.8	50	10	US-09-783-590-2759	Sequence 2759, Ap
c	30	18.8	0.8	42	10	US-09-904-380-8	Sequence 8, Appli
	31	18.8	0.8	47	9	US-09-853-526-283	Sequence 283, App
	32	18.8	0.8	47	10	US-09-901-484A-283	Sequence 283, App
	33	18.8	0.8	48	9	US-09-825-805-50	Sequence 50, Appl
c	34	18.6	0.8	33	10	US-09-971-611-25	Sequence 25, Appl
	35	18.6	0.8	45	9	US-10-043-415-1	Sequence 1, Appli
	36	18.6	0.8	48	9	US-09-864-785-3061	Sequence 3061, Ap
	37	18.6	0.8	48	9	US-09-864-785-3146	Sequence 3146, Ap
	38	18.6	0.8	48	9	US-09-864-785-3552	Sequence 3552, Ap
c	39	18.6	0.8	50	9	US-09-923-070A-13	Sequence 13, Appl
	40	18.4	0.7	48	9	US-09-864-785-3205	Sequence 3205, Ap
	41	18.4	0.7	48	9	US-09-864-785-3215	Sequence 3215, Ap
	42	18.4	0.7	48	9	US-09-864-785-3295	Sequence 3295, Ap
	43	18.4	0.7	48	9	US-09-864-785-3322	Sequence 3322, Ap
	44	18.4	0.7	48	9	US-09-864-785-3574	Sequence 3574, Ap
	45	18.4	0.7	48	9	US-09-864-785-3614	Sequence 3614, Ap

ALIGNMENTS

RESULT 1
US-09-426-548-106/c
; Sequence 106, Application US/09426548
; Patent No. US20010044936A1
; GENERAL INFORMATION:
; APPLICANT: Robbins, David
; APPLICANT: Lin-Goerke, Juili L.
; APPLICANT: Ling, Jessica
; TITLE OF INVENTION: No. US20010044936Alel Mutations in Human MLH1 and MSH2 Gene:
; TITLE OF INVENTION: Diagnosing Colorectal Cancer
; FILE REFERENCE: DEX-0054
; CURRENT APPLICATION NUMBER: US/09/426,548
; CURRENT FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-426-548-106

Query Match 0.9%; Score 22; DB 10; Length 46;
Best Local Similarity 73.7%; Pred. No. 5.9e+04;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy   1870 AAAGTAAATAAAAGTACAAAGCATATTTTAGTTAGTAC 1907
          |||  ||| |||||||||   || |||| || || |||
Db     43 AAAAAAAAAAAAAGTACATTACAGATTTGAGATATTAC 6
```

RESULT 2
US-09-842-883-13/c
; Sequence 13, Application US/09842883
; Patent No. US20020051770A1
; GENERAL INFORMATION:
; APPLICANT: Rovinski, Benjamin
; APPLICANT: Tartaglia, James
; APPLICANT: Cao, Shi-Xian
; APPLICANT: Persson, Roy
; APPLICANT: Klein, MIchel H.
; TITLE OF INVENTION: IMMUNIZING AGAINST HIV INFECTION
; FILE REFERENCE: 1038-1142 MIS

```
;  CURRENT APPLICATION NUMBER: US/09/842,883
;  CURRENT FILING DATE:  2001-04-27
;  PRICR APPLICATION NUMBER: 60/200,011
;  PRICR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus
US-09-842-883-13

  Query Match             0.9%;  Score 22;  DB 10;  Length 47;
  Best Local Similarity   73.7%;  Pred. No. 5.9e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    968 AGAGTGTCGTAGAAAGAAGAAAGAATATGTGAAATGTT 1005
          ||||||  | ||| |||| |||||||||    |||| |
Db     47 AGAGTGGTGCAGAGAGAAAAAAGAATATTCATAATGAT 10


RESULT 3
CS-09-842-883-14
; Sequence 14, Application US/09842883
; Patent No. US20020051770A1
; GENERAL INFORMATION:
;  APPLICANT: Rovinski, Benjamin
;  APPLICANT:  Tartaglia, James
;  APPLICANT:  Cao, Shi-Xian
;  APPLICANT:  Persson, Roy
;  APPLICANT:  Klein, MIchel H.
;  TITLE OF INVENTION: IMMUNIZING AGAINST HIV INFECTION
;  FILE REFERENCE: 1038-1142 MIS
;  CURRENT APPLICATION NUMBER: US/09/842,883
;  CURRENT FILING DATE:  2001-04-27
;  PRIOR APPLICATION NUMBER: 60/200,011
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus
US-09-842-883-14

  Query Match             0.9%;  Score 22;  DB 10;  Length 48;
  Best Local Similarity   73.7%;  Pred. No. 6e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    968 AGAGTGTCGTAGAAAGAAGAAAGAATATGTGAAATGTT 1005
          ||||||  | ||| |||| |||||||||    |||| |
Db      1 AGAGTGGTGCAGAGAGAAAAAAGAATATTCATAATGAT 38


RESULT 4
US-09-838-386-17
; Sequence 17, Application US/09838386
; Patent No. US20010055756A1
; GENERAL INFORMATION:
;  APPLICANT: Pellerin, Charles
;  APPLICANT:  Kukolj, George
;  TITLE OF INVENTION: Internal De No. US20010055756Alo Initiation Sites of the HCV NS5E
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 1011.2180001
;  CURRENT APPLICATION NUMBER: US/09/838,386
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 60/198,793
;  PRIOR FILING DATE: 2000-04-21
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: oligonucleotide
US-09-838-386-17

  Query Match             0.8%;  Score 20.8;  DB 10;  Length 45;
  Best Local Similarity   70.0%;  Pred. No. 1.1e+05;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1168 AATAAACATTTTATTTTCTAAACATTTCTTTTTTTCTATG 1207
          ||||  |  ||| |||| |     ||||||||||| || ||
Db      3 AATAGGCCATTTTTTTTTTTTTTTTTCTTTTTTCTTTG 42


RESULT 5
US-09-838-386-18/c
; Sequence 18, Application US/09838386
; Patent No. US20010055756A1
; GENERAL INFORMATION:
;  APPLICANT: Pellerin, Charles
;  APPLICANT:  Kukolj, George
;  TITLE OF INVENTION: Internal De No. US20010055756Alo Initiation Sites of the HC
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 1011.2180001
;  CURRENT APPLICATION NUMBER: US/09/838,386
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 60/198,793
;  PRIOR FILING DATE: 2000-04-21
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: oligonucleotide
US-09-838-386-18

  Query Match             0.8%;  Score 20.8;  DB 10;  Length 45;
  Best Local Similarity   70.0%;  Pred. No. 1.1e+05;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1168 AATAAACATTTTATTTTCTAAACATTTCTTTTTTTCTATG 1207
          ||||  |  ||| |||| |     ||||||||||||| ||
Db     43 AATAGGCCATTTTTTTTTTTTTTTTTCTTTTTTTCTTTG 4


RESULT 6
US-10-027-806-92/c
; Sequence 92, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
;  APPLICANT: Swanson, Ronald V.
;  APPLICANT:  Feldman, Robert A.
;  APPLICANT:  Schleper, Christa
;  TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
;  FILE REFERENCE: DCORP.002A
;  CURRENT APPLICATION NUMBER: US/10/027,806
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 92
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Cenarchaeum symbiosum
;   FEATURE:
;   NAME/KEY: TATA_signal
;   LOCATION: (11)...(16)
US-10-027-806-92
```

```
  Query Match             0.8%;  Score 20.4;  DB 9;  Length 43;
  Best Local Similarity   71.1%;  Pred. No. 1.3e+05;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      90 GGAGGAGCTTGTACCACCGGTAACTAAATGACCATGGA 127
           | || || ||||||| || |||| | |||||    |||
Db      39 GCAGCAGGTTGTACCCCCCGTAATTTAATGAGGCCGGA 2


RESULT 7
US-10-034-623-92/c
; Sequence 92, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
;  APPLICANT: Swanson, Ronald V.
;  APPLICANT:  Feldman, Robert A.
;  APPLICANT:  Schleper, Christa
;  TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
;  FILE REFERENCE: DCORP.002A
;  CURRENT APPLICATION NUMBER: US/10/034,623
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/408,020
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: 60/102,294
;  PRIOR FILING DATE: 1998-09-29
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 92
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Cenarchaeum symbiosum
;   FEATURE:
;   NAME/KEY: TATA_signal
;   LOCATION: (11)...(16)
US-10-034-623-92

  Query Match             0.8%;  Score 20.4;  DB 9;  Length 43;
  Best Local Similarity   71.1%;  Pred. No. 1.3e+05;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      90 GGAGGAGCTTGTACCACCGGTAACTAAATGACCATGGA 127
           | || || ||||||| || |||| | |||||    |||
Db      39 GCAGCAGGTTGTACCCCCCGTAATTTAATGAGGCCGGA 2


RESULT 8
US-09-853-526-206
; Sequence 206, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Ilya, Chumakov
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: PROSTATE CANCER GENE
;  FILE REFERENCE: GENSET.18CP1CP
;  CURRENT APPLICATION NUMBER: US/09/853,526
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: Patent.pm
; SEQ ID NO 206
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1..47
;   OTHER INFORMATION: polymorphic fragment 4-50-323
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: polymorphic base C
;   NAME/KEY: primer_bind
;   LOCATION: 1..23
;   OTHER INFORMATION: potential microsequencing oligo 4-50-323.mis1
;   NAME/KEY: primer_bind
;   LOCATION: 25..47
;   OTHER INFORMATION: complement potential microsequencing oligo 4-50-323.mis2
US-09-853-526-206

  Query Match             0.8%;  Score 20.4;  DB 9;  Length 47;
  Best Local Similarity   65.2%;  Pred. No. 1.3e+05;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    1060 TAAAAGCACTTAAGGACCTTTACTGCCACAAATCAGATTAATTTGG 1105
           | ||| || | | | | |||||||  | |||||    | |   |||| |
Db       2 TTAAAACATTGATGAATCTTTACTACTACAAAAGGGTTCGATTTAG 47


RESULT 9
US-09-901-484A-206
; Sequence 206, Application US/09901484A
; Patent No. US20020119460A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: Prostate Cancer Gene
;  FILE REFERENCE: GEN-T111XC3D2
;  CURRENT APPLICATION NUMBER: US/09/901,484A
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: US 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: US 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: US 09/853,526
;  PRIOR FILING DATE: 2001-05-11
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1)..(47)
;   OTHER INFORMATION: polymorphic fragment 4-50-323
;   NAME/KEY: allele
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: polymorphic base C
;   NAME/KEY: primer_bind
;   LOCATION: (1)..(23)
;   OTHER INFORMATION: potential microsequencing oligo 4-50-323.mis1
;   NAME/KEY: primer_bind
;   LOCATION: (25)..(47)
;   OTHER INFORMATION: complement potential microsequencing oligo 4-50-323.mis2
US-09-901-484A-206

  Query Match             0.8%;  Score 20.4;  DB 10;  Length 47;
  Best Local Similarity   65.2%;  Pred. No. 1.3e+05;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    1060 TAAAAGCACTTAAGGACCTTTACTGCCACAAATCAGATTAATTTGG 1105
```

```
Db      2 TTAAAACATTGATGAATCTTTACTACTACAAAAGGGTTCGATTTAG 47


RESULT 10
US-10-079-623-71/c
; Sequence 71, Application US/10079623
; Patent No. US20020169302A1
; GENERAL INFORMATION:
;  APPLICANT: Havukkala, Ilkka J.
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Grigor, Murray R.
;  APPLICANT:  Molenaar, Adrian J.
;  TITLE OF INVENTION: Compositions isolated from bovine
;  TITLE OF INVENTION:  mammary gland and methods for their use.
;  FILE REFERENCE: 11000.1044c3
;  CURRENT APPLICATION NUMBER: US/10/079,623
;  CURRENT FILING DATE:  2002-02-19
;  NUMBER OF SEQ ID NOS: 370
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Bovine
US-10-079-623-71

  Query Match             0.8%;  Score 20.2;  DB 9;  Length 44;
  Best Local Similarity   68.3%;  Pred. No. 1.4e+05;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1780 CAGTTCTGTATGTGTTCAACATTTTTGAATACATTAAAAGA 1820
Db     44 CAATTCTGGATATGCTTCACTTTGGTAAATATATTTTAAGA 4


RESULT 11
US-09-827-289-22
; Sequence 22, Application US/09827289
; Patent No. US20020009716A1
; GENERAL INFORMATION:
;  APPLICANT: Abarzua, Patricio
;  TITLE OF INVENTION: Process for Allele Discrimination Using Primer
;  TITLE OF INVENTION:  Extension
;  FILE REFERENCE: 469290-55
;  CURRENT APPLICATION NUMBER: US/09/827,289
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: U.S. 60/194843
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 46
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: P1 primer for
;   OTHER INFORMATION: use in allele discrimination
US-09-827-289-22

  Query Match             0.8%;  Score 20;  DB 10;  Length 46;
  Best Local Similarity   72.2%;  Pred. No. 1.6e+05;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1176 TTTTATTTTCTAAACATTTCTTTTTTTCTATGCGCA 1211
Db     11 TTTTTTTTTTTTTTTTTTTTTTTTTTCTATTCTCA 46


RESULT 12
US-09-864-785-3523
; Sequence 3523, Application US/09864785
; Patent No. US20020177568A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Stinchcomb, Dan
;  APPLICANT:  Draper, Ken
;  APPLICANT:  McSwiggen, Jim
;  TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions F
;  TITLE OF INVENTION:  Levels of NF-Kappa B
;  FILE REFERENCE: 400/022 (MBHB00-812-D)
;  CURRENT APPLICATION NUMBER: US/09/864,785
;  CURRENT FILING DATE:  2001-05-23
;  NUMBER OF SEQ ID NOS: 3929
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3523
;   LENGTH: 48
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Nucleic Acid
US-09-864-785-3523

  Query Match             0.8%;  Score 20;  DB 9;  Length 48;
  Best Local Similarity   61.1%;  Pred. No. 1.6e+05;
  Matches   22;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    438 CAAAAGCGAAGGGAAATTCTTTCAAGGAGGCCTTCC 473
Db      2 CAAAGGCGGAGGAAACUCCCUUCAAGGACAUCGUCC 37


RESULT 13
US-09-729-402-28/c
; Sequence 28, Application US/09729402
; Patent No. US20010021379A1
;   GENERAL INFORMATION:
;        APPLICANT: Cousens, Lawrence S.
;                   Eberhardt, Christine D.
;                   Gray, Patrick W.
;                   Le Trong, Hai
;                   Tjoelker, Larry W.
;                   Wilder, Cheryl L.
;        TITLE OF INVENTION: Platelet-Activating Factor
;                            Acetylhydrolase
;        NUMBER OF SEQUENCES: 30
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/729,402
;             FILING DATE: 04-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/318,905
;             FILING DATE: 06-OCT-1994
;             APPLICATION NUMBER: US 08/133,803
;             FILING DATE: 06-OCT-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: No. US20010021379Aland, Greta E.
;             REGISTRATION NUMBER: 35,302
;             REFERENCE/DOCKET NUMBER: 27866/32793
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (312) 474-6300
;             TELEFAX: (312) 474-0448
;             TELEX: 25-3658
;   INFORMATION FOR SEQ ID NO: 28:
;        SEQUENCE CHARACTERISTICS:
```

```
;           LENGTH: 41 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: single
;           TOPOLOGY: linear
;       MOLECULE TYPE: DNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-729-402-28

  Query Match             0.8%;  Score 19.8;  DB 10;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.5e+05;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    707 TGCAGCAGCCACTCAGCCGGGTACTACCATTCTACAGTA 745
          ||||||||||| | |  |  |||  |  |||||| | ||
Db     39 TGCAGCAGCCATTAATACTTGTATCATAATTCTAGAATA 1


RESULT 14
US-09-921-398-18
; Sequence 18, Application US/09921398
; Patent No. US20020055169A1
;   GENERAL INFORMATION:
;       APPLICANT: Tekamp-Olson, Patricia
;       TITLE OF INVENTION: METHOD FOR EXPRESSION OF HETEROLOGOUS
;                           PROTEINS IN YEAST
;       NUMBER OF SEQUENCES: 41
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Bell Seltzer IP Group of Alston & Bird, LLP
;           STREET: 3605 Glenwood Ave. Suite 310
;           CITY: Raleigh
;           STATE: NC
;           COUNTRY: US
;           ZIP: 27622
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/921,398
;           FILING DATE: 02-Aug-2001
;           CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Spruill, W. Murray
;           REGISTRATION NUMBER: 32,943
;           REFERENCE/DOCKET NUMBER: 5784-4
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 919 420 2202
;           TELEFAX: 919 881 3175
;   INFORMATION FOR SEQ ID NO: 18:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 45 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: single
;           TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       ORIGINAL SOURCE:
;           ORGANISM: Homo sapiens
;       SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-921-398-18

  Query Match             0.8%;  Score 19.8;  DB 10;  Length 45;
  Best Local Similarity   69.2%;  Pred. No. 1.7e+05;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    798 GTTCAAGCTGCCTCTGGAGACGTACAAACATACCAGATT 836
          ||||||| |  ||||||  | || ||| | | | || ||
Db      2 GTTCAAGTTAACTCTGGTCAAGTTCAATCTTTCTAGCTT 40


RESULT 15
US-09-263-959-486
; Sequence 486, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  486:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-486

  Query Match             0.8%;  Score 19.8;  DB 10;  Length 45;
  Best Local Similarity   77.4%;  Pred. No. 1.7e+05;
  Matches   24;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   1176 TTTTATTTTCTAAACATTTCTTTTTTTCTAT 1206
          |||| ||||||   | ||||||||||| | |
Db      2 TTTTCTTTTCTTTTCCTTTCTTTTTTTTTT 32


Search completed: March  9, 2003, 21:49:51
Job time : 199 secs
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 18:54:42 ; Search time 3109 Seconds
(without alignments)
12783.443 Million cell updates/sec

Title: US-09-973-827-3
Perfect score: 2454
Sequence: 1 cgcggcggaggtgtagtttg.....tctgtatttgtgtatatcat 2454

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 102860

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	Length	DB	ID	Description
	1	44	1.8	50	9	AU107544	AU107544 AU107544
	2	22.8	0.9	50	17	AZ663650	AZ663650 1M0543G23
c	3	22.6	0.9	46	9	AU014208	AU014208 AU014208
c	4	22.4	0.9	48	17	AZ361863	AZ361863 1M0106B17
c	5	22.2	0.9	46	9	AU014238	AU014238 AU014238
c	6	22	0.9	50	9	AA545213	AA545213 vj93e11.r
c	7	22	0.9	50	13	BI425819	BI425819 sah72g08.
c	8	21.8	0.9	44	17	AL757115	AL757115 Arabidops
c	9	21.8	0.9	46	17	AZ820550	AZ820550 2M0092M18
	10	21.6	0.9	48	13	BJ083775	BJ083775 BJ083775
c	11	21.4	0.9	46	10	AV851679	AV851679 AV851679
	12	21.2	0.9	43	17	AZ638328	AZ638328 1M0498I10
c	13	21.2	0.9	45	9	AU014228	AU014228 AU014228
	14	21	0.9	46	10	AV947975	AV947975 AV947975
	15	21	0.9	50	10	AW333744	AW333744 S25E9 AGS
c	16	20.8	0.8	46	17	AL760388	AL760388 Arabidops
	17	20.8	0.8	50	9	AU103894	AU103894 AU103894
	18	20.6	0.8	49	17	BH861575	BH861575 SALK_0875
	19	20.4	0.8	43	17	AZ439166	AZ439166 1M0229J07
c	20	20.4	0.8	45	2	HSM002460	A1038121 Homo sapi
c	21	20.4	0.8	46	13	BJ015738	BJ015738 BJ015738
c	22	20.4	0.8	46	13	BJ075979	BJ075979 BJ075979
c	23	20.4	0.8	50	9	AJ499559	AJ499559 AJ499559
	24	20.4	0.8	50	13	BJ066224	BJ066224 BJ066224
	25	20.2	0.8	45	9	AU269673	AU269673 AU269673
c	26	20.2	0.8	45	17	AL761682	AL761682 Arabidops
	27	20.2	0.8	48	17	HSCHO6F06	X88550 H.sapiens D
c	28	20.2	0.8	50	17	BH791446	BH791446 SALK_0599
	29	20	0.8	43	9	AA509550	AA509550 vh50d03.r
	30	20	0.8	43	17	AZ620145	AZ620145 1M0452D15
	31	20	0.8	44	14	H98532	H98532 yv97d03.s1
	32	20	0.8	45	17	AZ427755	AZ427755 1M0209N24
	33	20	0.8	46	17	AZ766403	AZ766403 1M0563O23
	34	20	0.8	48	17	AZ592310	AZ592310 1M0403D12
	35	20	0.8	49	17	TA98D04Q	AL459661 T. brucei
c	36	20	0.8	50	9	AU103398	AU103398 AU103398
c	37	19.8	0.8	32	17	AL761349	AL761349 Arabidops
	38	19.8	0.8	40	17	AQ073791	AQ073791 EP(3)3174
c	39	19.8	0.8	40	17	TA105A11P	AL464696 T. brucei
c	40	19.8	0.8	46	9	AA946759	AA946759 oq50c05.s
	41	19.8	0.8	46	10	AW334262	AW334262 S32E7 AGS
c	42	19.8	0.8	46	13	BI256739	BI256739 602974296
	43	19.8	0.8	46	17	AL770028	AL770028 Arabidops
c	44	19.8	0.8	47	17	AZ843462	AZ843462 2M0142P18
c	45	19.8	0.8	48	2	HSM002426	A1038087 Homo sapi

ALIGNMENTS

RESULT 1
AU107544
LOCUS AU107544 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU107544 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone NBLAN589NF, mRNA sequence.
ACCESSION AU107544
VERSION AU107544.1 GI:13557065
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers

```
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NBLAN589NF"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        8 a      6 c     25 g     11 t
ORIGIN
```

Query Match 1.6%; Score 44; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CGCGGCGGAGGTGTAGTTTGACGCGGTGTGTTACGTGGGGGAGA 44
          ||||||||||||||||||||||||||||||||||||||||||||
Db       7 CGCGGCGGAGGTGTAGTTTGACGCGGTGTGTTACGTGGGGGAGA 50
```

```
RESULT 2
AZ663650
LOCUS       AZ663650                 50 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0543G23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0543G23 F, DNA sequence.
ACCESSION   AZ663650
VERSION     AZ663650.1  GI:11800796
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0543  row: G  column: 23
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 50.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0543G23"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       15 a     17 c      5 g     13 t
ORIGIN
```

Query Match 0.9%; Score 22.8; DB 17; Length 50;
Best Local Similarity 71.4%; Pred. No. 1.6e+06;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy     795 GTTGTTCAAGCTGCCTCTGGAGACGTACAAACATACCAGATT 836
           ||  || |  || |  ||||| || |||| |||||| |||||
Db       1 GTCTTTGACACTACTCCTGGACACCTACACACATACAAGATT 42
```

```
RESULT 3
AU014208/c
LOCUS       AU014208                 46 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU014208 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc09381, mRNA sequence.
ACCESSION   AU014208
VERSION     AU014208.1  GI:3368999
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc09381"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       22 a      5 c      3 g     15 t      1 others
ORIGIN
```

Query Match 0.9%; Score 22.6; DB 9; Length 46;
Best Local Similarity 67.4%; Pred. No. 1.8e+06;
Matches 31; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy    2122 TTTATGAAATGTTACATGTAGAAAAATACTGATTTTAAATATTTTC 2167
           ||| |  | ||| | | |||| |  ||| ||||||||| | || | |
Db      46 TTTTTTCACTGTAAAAGGTAGTATTATATTGATTTTATAAAATTNC 1
```

```
RESULT 4
AZ361863/c
LOCUS       AZ361863                 48 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0106B17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0106B17 R, DNA sequence.
```

```
ACCESSION   AZ361863
VERSION     AZ361863.1  GI:10475563
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0106  row: B  column: 17
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0106B17"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       10 a      13 c       8 g      17 t
ORIGIN

  Query Match           0.9%;  Score 22.4;  DB 17;  Length 48;
  Best Local Similarity   72.5%;  Pred. No. 2e+06;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    1007 AGAAAACAGAGTGGCAGTGCTTGAAAATCAAAACAAGACA 1046
           ||| || |||| ||||||  ||| | ||| ||  ||| ||
Db      42 AGATAATAGAGGGGCAGTCTTTGTACATCCAAGTAAGCCA 3


RESULT 5
AU014238/c
LOCUS       AU014238                 46 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU014238 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc09421, mRNA sequence.
ACCESSION   AU014238
VERSION     AU014238.1  GI:3369029
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc09421"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       22 a       5 c       3 g      13 t       3 others
ORIGIN

  Query Match           0.9%;  Score 22.2;  DB 9;  Length 46;
  Best Local Similarity   65.2%;  Pred. No. 2.2e+06;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    2122 TTTATGAAATGTTACATGTAGAAAAATACTGATTTTAAATATTTTC 2167
           |||.|  | ||| | | |||| |  ||| |||||||:|   | || |
Db      46 TTTTTTCACTGTAANAGGTAGTATTATATTGATTTTATNAAATTNC 1


RESULT 6
AA545213/c
LOCUS       AA545213                 50 bp    mRNA    linear   EST 04-AUG-1997
DEFINITION  vj93e11.r1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
            IMAGE:944684 3', mRNA sequence.
ACCESSION   AA545213
VERSION     AA545213.1  GI:2306287
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
            MGI:541540.
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="Mus musculus"
                      /strain="C57BL/6J x DBA/2J F1"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:944684"
                      /clone_lib="Knowles Solter mouse 2 cell"
                      /tissue_type="embryo"
                      /dev_stage="2-cell"
                      /lab_host="DH10B"
                      /note="Organ: embryo; Vector: pBluescribe (modified);
                      Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                      mRNA prepared from 13,500 2-cell stage embryos. Primer:
                      SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'. cDNAs
                      were cloned into the MluI/SalI sites of a modified
                      pBluescribe vector using commercial linkers (NEB).
                      Average insert size: 1.2 kb."
BASE COUNT       29 a      5 c      7 g      9 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 9;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 2.4e+06;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    2330 TTTTTTTGTTCTCTTGTAAAAAGAGTAGTTATTAGTTCTGCTTTAG 2375
           ||||||| || | ||| ||| || || || ||   | | ||||
Db      50 TTTTTTTTTTTTTTTGGAAACTGATTATTTCTTCTGTTTACCTTAG 5


RESULT 7
BI425819/c
LOCUS       BI425819                 50 bp    mRNA    linear   EST 29-NOV-2001
DEFINITION  sah72g08.y1 Gm-c1049 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1049-4143 5', mRNA sequence.
ACCESSION   BI425819
VERSION     BI425819.1  GI:15203051
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Putative full length read
            vector to vector length is 51 This clone is available through:
            ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
            35801 For further information call: (800)-533-4363 or contact via
            email: ccu@resgen.com.
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="Glycine max"
                      /db_xref="taxon:3847"
                      /clone="GENOME SYSTEMS CLONE ID: Gm-c1049-4143"
                      /clone_lib="Gm-c1049"
                      /tissue_type="whole seedlings of greenhouse grown plants"
                      /dev_stage="3 week old"
                      /lab_host="DH10B"
                      /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                      XhoI; The Clark NIL was constructed and seed was provided
                      by Dr. J. Specht, University of Nebraska (Shoemaker and
                      Specht, 1995). The cDNA library was constructed from mRNA
                      isolated from whole seedlings of 3 week old greenhouse
                      grown plants. Complementary DNA was synthesized from mRNA
                      using a primer consisting of a poly(dT) sequence with a
                      XhoI restriction site and a 3' anchor. EcoRI adapters were
                      ligated to the blunt-ended cDNA fragments followed by XhoI
                      digestion. The cDNA fragments were directionally cloned
                      into the EcoRI-XhoI restriction site of the pBluescript
                      vector. The ligated cDNA fragments were transformed into
                      DH10B host cells (GibcoBRL). The library was constructed
                      in cooperation with Dr. Paul Keim's laboratory at Northern
                      Arizona University."
BASE COUNT       10 a      7 c      7 g      26 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 13;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 2.4e+06;
  Matches   25;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    1400 AGAAATGAGGAAAAGAAAATCTTTTAAAA 1429
           ||||| ||||||||||| |||| || |||
Db      40 AGAAAAAAGGAAAAGAAATTCTTATTTAAA 11


RESULT 8
AL757115/c
LOCUS       AL757115                 44 bp    DNA     linear   GSS 18-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-118A06-012518,
            genomic survey sequence.
ACCESSION   AL757115
VERSION     AL757115.1  GI:21495463
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 44)
  AUTHORS   Rosso,M., Strizhov,N., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone T11I11.
            The sequences are generated at the MPI for Plant Breeding Research
            in the context of the GABI-Kat project. GABI-Kat is part of the
            German Plant Genomics program designated 'GABI'. Information on
            line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES              Location/Qualifiers
     source           1. .44
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="GK-118A06-012518"
                      /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                      /note="PCR was performed on DNA from Arabidopsis thaliana
```

plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 10 a 13 c 3 g 18 t
ORIGIN

Query Match 0.9%; Score 21.8; DB 17; Length 44;
Best Local Similarity 70.7%; Pred. No. 2.7e+06;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1516 CATGCATAAAGTAAGTAAGGTGCAATGAAGAAGTGTTGATT 1556
Db 41 CATAGATACAGAGAGAGAGATAGAATGAAGGAGTGTAGATT 1

RESULT 9
AZ820550/c
LOCUS AZ820550 46 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0092M18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0092M18 R, DNA sequence.
ACCESSION AZ820550
VERSION AZ820550.1 GI:12990458
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0092 row: M column: 18
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 46.
FEATURES Location/Qualifiers
source 1. .46
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0092M18"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 8 c 3 g 21 t
ORIGIN

Query Match 0.9%; Score 21.8; DB 17; Length 46;
Best Local Similarity 78.8%; Pred. No. 2.7e+06;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2203 TTTTTAAATGAAAGAAGTTGTAAGGATATAAAA 2235
Db 45 TGTATGAATGCATGAATTTATAAGGATATAAAA 13

RESULT 10
BJ083775
LOCUS BJ083775 48 bp mRNA linear EST 12-DEC-2001
DEFINITION BJ083775 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL086c24 3', mRNA sequence.
ACCESSION BJ083775
VERSION BJ083775.1 GI:17578994
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 48)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .48
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL086c24"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute). "

BASE COUNT 27 a 0 c 4 g 16 t 1 others
ORIGIN

Query Match 0.9%; Score 21.6; DB 13; Length 48;
Best Local Similarity 66.7%; Pred. No. 2.9e+06;
Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1160 AAAGACAAAATAAACATTTTATTTTCTAAACATTTCTTTTTTTCT 1204
Db 4 AAAAAAAAAAAAAAAAAAAAAATTTTTAAAAAGGGNTTTTTTTTT 48

RESULT 11
AV851679/c

```
LOCUS       AV851679                 46 bp    mRNA    linear   EST 08-NOV-2001
DEFINITION  AV851679 Nori Satoh unpublished cDNA library, larva Ciona
            intestinalis cDNA clone rcilv05b16 3', mRNA sequence.
ACCESSION   AV851679
VERSION     AV851679.1  GI:16836144
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="rcilv05b16"
                     /clone_lib="Nori Satoh unpublished cDNA library, larva"
                     /tissue_type="whole animal"
                     /dev_stage="larva"
                     /note="Vector: pBluescript SK"
BASE COUNT       18 a     11 c      5 g     11 t      1 others
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 10;  Length 46;
  Best Local Similarity   70.0%;  Pred. No. 3.2e+06;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    2087 CAAGATGTCAAATAAAGCAAAGTGATATATATTTGTTTAT 2126
           || ||||||   | | ||| || |||| |  |||| ||||
Db      45 CATGATGTCNTTTTACGCAGAGGGATACAAGTTTGGTTAT 6


RESULT 12
AZ638328
LOCUS       AZ638328                 43 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0498I10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0498I10 F, DNA sequence.
ACCESSION   AZ638328
VERSION     AZ638328.1  GI:11760518
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0498  row: I  column: 10
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 43.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0498I10"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        6 a      1 c      1 g     35 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 17;  Length 43;
  Best Local Similarity   69.0%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    1163 GACAAAATAAACATTTTATTTTCTAAACATTTCTTTTTTTCT 1204
           |||||||      |||| |||| | |   ||| |||||||
Db       1 GACAAAATCTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTT 42


RESULT 13
AU014228/c
LOCUS       AU014228                 45 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU014228 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc09407, mRNA sequence.
ACCESSION   AU014228
VERSION     AU014228.1  GI:3369019
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc09407"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
```

```
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       21 a       5 c       3 g      14 t       2 others
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 9;  Length 45;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    2122 TTTATGAAATGTTACATGTAGAAAAATACTGATTTTAAATATT 2164
           ||| |  | ||| | | |||| |  ||| |||||||| | | |
Db      45 TTTTTTCACTGTAAAAGGTAGTATTATATTGATTTTATANAAT 3


RESULT 14
AV947975
LOCUS       AV947975                 46 bp    mRNA    linear   EST 14-MAR-2002
DEFINITION  AV947975 Nori Satoh unpublished cDNA library, young adult Ciona
            intestinalis cDNA clone ciad05102 5', mRNA sequence.
ACCESSION   AV947975
VERSION     AV947975.1  GI:19425734
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Satoh,N., Satou,Y., Kchara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="ciad05102"
                     /clone_lib="Nori Satoh unpublished cDNA library, young
                     adult"
                     /tissue_type="whole animal"
                     /dev_stage="young adult"
                     /note="Vector: pBluescript SK"
BASE COUNT       18 a       0 c       4 g      22 t       2 others
ORIGIN

  Query Match             0.9%;  Score 21;  DB 10;  Length 46;
  Best Local Similarity   69.2%;  Pred. No. 3.9e+06;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    1176 TTTTATTTTCTAAACATTTCTTTTTTTCTATGCGCAAAA 1214
           |||| ||||  ||| |  | |||||||  |  | | ||||
Db       2 TTTTTTTTTAAAAAAAANTTTTTTTTTTTTNGGGGAAAA 40


RESULT 15
AW333744
LOCUS       AW333744                 50 bp    mRNA    linear   EST 31-JAN-2000
DEFINITION  S25E9 AGS-1 Pneumocystis carinii f. sp. carinii cDNA 3', mRNA
            sequence.
ACCESSION   AW333744
VERSION     AW333744.1  GI:6830101
KEYWORDS    EST.
SOURCE      Pneumocystis carinii f. sp. carinii.
  ORGANISM  Pneumocystis carinii f. sp. carinii
            Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
            Pneumocystidaceae; Pneumocystis.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Smulian,A.G., Arnold,J., Weise,M., Wunderlich,J., Staben,C., Edman
            ,J.C., Kovacs,J. and Cushion,M.
  TITLE     Expressed sequence tags from Pneumocystis carinii
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Staben C
            School of Biological Sciences
            University of Kentucky
            101 Morgan Building, University of Kentucky, Lexington, KY
            40506-0225, USA
            Tel: 606 257 2161
            Fax: 606 257 1717
            Email: staben@pop.uky.edu.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Pneumocystis carinii f. sp. carinii"
                     /db_xref="taxon:38081"
                     /clone_lib="AGS-1"
                     /lab_host="E. coli"
                     /note="Vector: Lambda ZAP II; Site_1: EcoRI; Site_2: XhoI;
                     P. carinii organisms (3x10e9) from a single rat (99-1-6,
                     sacrificed on 3/17/99)  at Cincinnati VA facilities.
                     Trizol extracted RNA.  Oligo dT priming, standard
                     conditions described by  vendor, Stratagene. Further
                     details see www.uky.edu/Project/Pneumocystis/"
BASE COUNT       21 a       1 c       1 g      27 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 10;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 3.9e+06;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    2330 TTTTTTTGTTCTCTTGTAAAAAGAGTAGTTATTAGTTCTGCTTTA 2374
           ||||||| || | ||  |||| ||  |  |||||  |||  || |
Db       4 TTTTTTTTTTTTTTTAAAAAATGAAAAAATATTAAATCTATTTAA 48


Search completed: March  9, 2003, 21:44:26
Job time : 3114 secs
```